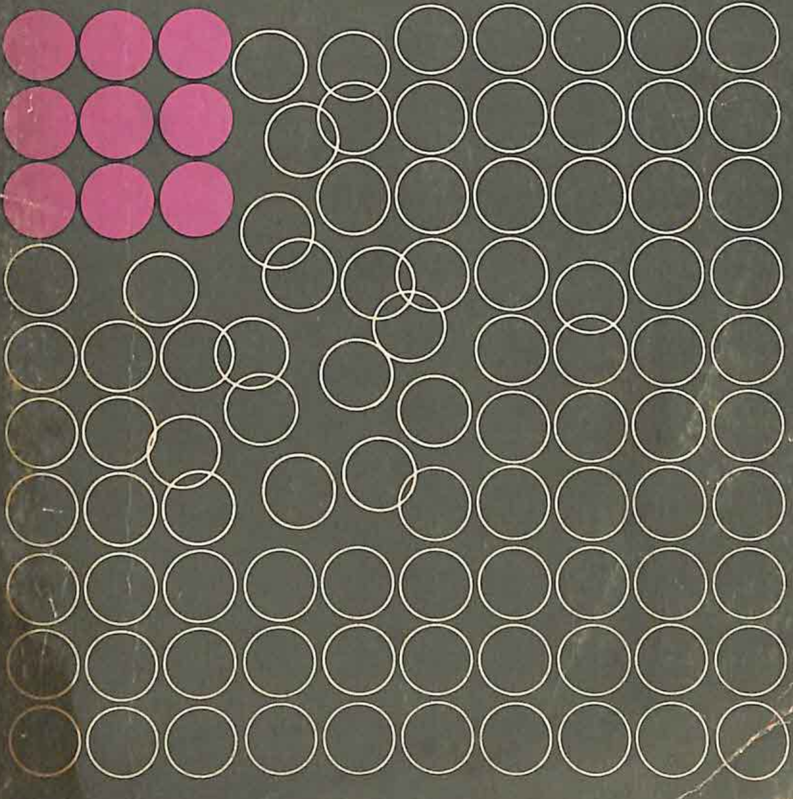
Penguin science of behaviour

Assessment in clinical psychology

C. E. Gathercole



2147
9.1.70

Penguin Education X14
Penguin Science of Behaviour
General Editor: B. M. Foss

Clinical Psychology
Editor: Graham A. Foulds

Assessment in Clinical Psychology

C. E. Gathercole

Penguin Science of Behaviour
General Editor: B. M. Foss
Professor of Psychology, Bedford College,
University of London

Abnormal Psychology
Editor: Max Hamilton
Nuffield Professor of Psychiatry,
University of Leeds

Attention and Set
Editor: Graham A. Foulds
Institute of Experimental
Psychology, University of
Oxford

Clinical Psychology
Editor: Anne Treisman
University of Edinburgh

Developmental Psychology
Editor: B. M. Foss
Professor of Psychology,
Bedford College,
University of London

Method and History
Editor: W. M. O'Neil
Deputy Vice-Chancellor,
University of Sydney

Motivation and Emotion
Editors: Dalbir Bindra
Professor of Psychology,
McGill University, Montreal
Jane Stewart
Associate Professor of
Psychology, Sir George
Williams University, Montreal

Perception and Thinking
Editor: P. C. Dodwell
Professor of Psychology,
Queen's University at
Kingston, Ontario

Physiological Psychology
Editor: K. H. Pribram
Research Professor of
Psychology and Psychiatry,
Stanford University

Skills and Learning
Editors: Harry Kay
Professor of Psychology,
University of Sheffield
L. J. Postman
Professor of Psychology,
Institute of Human Learning,
University of California,
Berkeley

Social Psychology
Editor: Michael Argyle
Institute of Experimental
Psychology, University of
Oxford

Assessment in Clinical Psychology

C. E. Gathercole

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

Penguin Books

Penguin Books Ltd, Harmondsworth,
Middlesex, England
Penguin Books Inc., 7110 Ambassador Road,
Baltimore, Md 21207, U.S.A.
Penguin Books Australia Ltd, Ringwood,
Victoria, Australia

First published 1968
Copyright © C. E. Gathercole, 1968

Made and printed in Great Britain by
Hazell Watson & Viney Ltd,
Aylesbury, Bucks
Set in Linotype Plantin

132.075
GAT

S.C.E R T., West Bengal
Date 9...1...70..........
Acc. No. 2147..................

This book is sold subject to the conditions that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

Penguin Science of Behaviour

This book is one of the first in an ambitious project, the *Penguin Science of Behaviour*, which will cover a very wide range of psychological inquiry. Many of the short 'unit' texts will be on central teaching topics, while others will deal with present theoretical and empirical work which the Editors consider to be important new contributions to psychology. We have kept in mind both the teaching divisions of psychology and also the needs of psychologists at work. For readers working with children, for example, some of the units in the field of Developmental Psychology will deal with techniques in testing children, other units will deal with work on cognitive growth. For academic psychologists, there will be units in well-established areas such as Learning and Perception, but also units which do not fall neatly under any one heading, or which are thought of as 'applied', but which nevertheless are highly relevant to psychology as a whole.

The project is published in short units for two main reasons. Firstly, a large range of short texts at inexpensive prices gives the teacher a flexibility in planning his course and recommending texts for it. Secondly, the pace at which important new work is published requires the project to be adaptable. Our plan allows a unit to be revised or a fresh unit to be added with maximum speed and minimal cost to the reader.

Above all, for students, the different viewpoints of many authors, sometimes overlapping, sometimes in contradiction, and the range of topics Editors have selected will reveal the complexity and diversity which exist beyond the necessarily conventional headings of an introductory course.

B.M.F.

For Dorrie

Contents

Editorial Foreword 11

Preface 15

1 Introduction 17

2 Assessment Referrals 29

3 Methods of Assessment 44

4 Conceptualizing the Patient 61

5 Selecting Tests 83

6 Base Rates and the Usefulness of Tests 93

7 Assessment Technique 109

8 Research in Assessment 133

Appendix 1 Descriptions of Selected Tests 144

Appendix 2 Tests Used in Clinical Psychology 154

References 161

Index 168

Editorial Foreword

This is the first book to be presented in the *Penguin Science of Behaviour* project in the field of Clinical Psychology. It offers an over-view of the types of activity in which the clinical psychologist engages before turning to the rather specific areas of interest constituted by the later texts on the various process disorders. The author accordingly describes some of the problems with which psychiatrists and others present the clinical psychologist. These problems include diagnostic assessment, evaluation of the effects of treatment, quantification of personality characteristics, determination of intellectual level, speed and persistence and their appropriateness for a particular treatment or occupation, and many others. The validity and reliability of techniques used to acquire data relevant to these problems are discussed, as are the methods the clinical psychologist adopts for interpreting and reporting on such data.

Like very many of his colleagues, the author is both a research worker and a practising clinician. He shows himself well aware of the desirability of avoiding unnecessary dissociation of the two roles. A gap does, and perhaps always will exist, so that the clinical psychologist has to develop a high tolerance for ambiguity. He has to be prepared to guess, and to make explicit the occasions on which he is doing so. The author has tried, and succeeding authors in this series will try also, to provide data which can make these guesses as informed as possible.

I have hinted that the series is presented under the disorders of process and a few notes on the structure of the series may be helpful in this introductory volume. In planning a group of books on clinical psychology, the main choice with which one is presented is between grouping by diagnosis or by process. Probably the majority of clinical psychologists believe that the classification of psychiatric illnesses is neither sufficiently reliable nor sufficiently free from arbitrariness to serve our purpose. When, however, one tries to group in terms of processes – attending, remembering, perceiving, thinking and so forth – it immediately becomes apparent that these distinctions are no less unreliable or arbitrary. If one opts for diagnostic grouping, processes of no known striking relevance to a particular illness may tend to be neglected; if one opts for process grouping, illnesses of no known striking relevance to the particular process may tend to be neglected. In either case the subsidiary mode of classification would be lost. In the absence of compelling reasons for rejecting one rather than the other, it was decided to retain both.

The primary grouping is in terms of disorders of processes, the normal functioning of which will often be treated more fully in books in other fields of this *Penguin Science of Behaviour* project. There will, accordingly, in the Clinical Psychology section, be books on disorders of attention, of memory, of perception, of thinking, etc. It is not, of course, possible to write about perception without frequent reference to attention or motivation for example. The book title indicates the process on which the particular author will focus his attention. Processes which will be peripheral for him, will be focal in other books. These overlaps present great difficulties for individual authors, but it is neither realistic nor desirable to try to eliminate them.

The secondary mode of classification is the diagnostic one and takes place inside each book. Here it is hoped that at least the main diagnostic classes will be adhered to – the personality disorders, the neuroses, the functional and the organic psychoses. Any reader who wishes to learn what clinical psychologists have to say, for example, about the neuroses will be able, as it were, to turn the classification upside down and to work through the unit texts within the framework: the neuroses – disorders of attention, disorders of learning and memory, etc. When there is little or nothing to say under any particular heading, authors will point this out so that the reader will be able to rest assured that the area has not been overlooked.

A subsidiary mode of classification within the primary mode will often be: identification, assessment and measurement of the disordered process; conditions under which the process disorder may arise; conditions under which it may be modified and conditions under which it might be prevented from arising.

Clinical psychology, which is virtually a post-war development, has been most closely associated with psychiatry. As Mr Gathercole points out, one expects rapid change in such a new discipline and he discusses the opportunities presented by automation and increasing computer facilities. Another development may well be in the direction of increasing contact with other medical specialities and with sociology, and it will be surprising if, in the next decade, there is not much to report in these and other areas.

G.A.F.

Preface

The aim of this book is to provide an introduction and a framework to some of the main issues of assessment. There are three kinds of student for which the book, though it makes no claims to be comprehensive, may be useful. First, undergraduate psychology students may find it a useful adjunct to courses in personality assessment. Next, for psychiatrists in training, it offers an outline of the approaches to assessment used by psychologists. Effective mutual collaboration between psychologists and psychiatrists can only be improved as psychiatrists increase their understanding of these approaches. Finally, and most directly, to the postgraduate student of clinical psychology who needs to orientate himself to the problems he will face in the practice of assessment.

The bulk of the work of clinical psychologists continues to be done in conjunction with psychiatrists, although a growing number are working in various other settings, such as in consultation with other medical specialists in general hospitals and with general practitioners and in private practice. This book is concerned with the context of adult psychiatry so the special problems of work in the psychiatric sub-specialties, such as forensic psychiatry, child psychiatry and mental subnormality, have been given little attention here, although the main issues raised have application in those fields.

An attempt has been made to cover some topics which receive little attention in the literature but which have a vigorous oral tradition among psychologists. The chapter on referral questions is a case in point. The emphasis of the book is on practical issues which frequently bewilder the student in his first contact with clinical psychology and for which little guidance is available in the larger textbooks. Where theoretical and conceptual issues are discussed the intention is to alert the student to topics he will encounter in more advanced reading.

In such a new field as clinical psychology one expects to find variations in outlook and practice. One also expects to find rapid change and development in various directions. I have attempted to present this variety without being too partisan.

I am indebted to the many psychologists and psychiatrists who have contributed to this book by their sympathetic and constructive discussions. Thanks are due to Mr R. J. McGuire who read the manuscript and made many useful suggestions, to Dr T. Freeman and Professor F. Fish for their encouragement and to my father E. C. Gathercole for his editorial work.

1 Introduction

The Need for Psychological Tests

The purpose of assessing a psychiatric patient is to provide information on which decisions about his future can be based. What treatment should he have? Should he be taken into hospital? Should he be discharged? Should he be persuaded not to take up his former employment? Should he have further investigations? These are some of the practical day-to-day decisions a psychiatrist has to make.

In order to make these decisions he requires information about the patient, some of which he will gather from his examination of the patient's mental state. He evaluates his findings in the light of the patient's history, family history and his physical condition and then formulates a tentative opinion. Often further information will be required from various sources so that a more precise reformulation can be made. Then decisions regarding the patient's treatment and disposal will be made. Such decisions involve prediction. For instance a patient with somatic complaints resulting from tension is believed to have numerous problems which he avoids facing. It is predicted that if he is confronted with his problems and manages to sort them out his tension and consequently his somatic complaints also will be reduced.

The psychiatrist's technique is to scan the information available to him for significant features which will enable him to account for the patient's symptoms. He considers various possible explanations and gradually by a process of elimination narrows down the possibilities. Since he is looking for features which can be followed up later, these scanning techniques are necessarily brief. As a result they are relatively crude and insensitive so that only the grosser forms of deficit and disorder will be caught in the net (see the discussion on bandwidth and fidelity, chapter 5). In addition, these procedures are very subjective. In one study, a correlation of only 0·25 was found between experienced psychiatrists rating anxiety in 80 psychiatric patients (Cattell, 1957). Often inferences about the patient are made which are greatly influenced by irrelevancies. The impressive appearance or cultured accent of a patient may influence a judgement about his intelligence. The personality of the psychiatrist may influence his conclusions regarding patients through counter-transference and projection. One psychiatrist sees hostility in the eyes of every patient whereas another cannot see hostility at all.

The need for developing psychological techniques to assess the behaviour of psychiatric patients has long been recognized by psychiatrists. Kraepelin studied with Wundt and used tests of memory, fatigue and learning ability in his psychiatric research. Jung experimented with word association tests and the galvanic skin response (G.S.R.). (See Appendix 1 for notes on tests.) Kretschmer made considerable use of psychological tests in his psychiatric research.

It is interesting that a number of tests which are now used mainly by clinical psychologists were introduced by psychiatrists: for example Rorschach's inkblots (1942) and Lauretta Bender's Visual Motor Gestalt Test (1938). Several tests have resulted from collaboration between psychiatrists and psychologists: for example the Minnesota Multiphasic Personality Inventory (M.M.P.I.) (Hathaway and McKinley, 1951), Progressive Matrices (Penrose and Raven, 1937), Tests of Abstract and Concrete Thinking (Goldstein and Scheerer, 1941), Family Relations Indicator (Howells and Lickorish, 1962). Although psychiatrists often employ psychological tests themselves, particularly in the examination of organic patients (Klein and Mayer-Gross, 1957), testing has come to be regarded mainly as the province of psychologists.

This is probably because research into tests has been done mainly by psychologists who have thereby acquired expertise in test development.

The training of the psychiatrist is centred on abnormal behaviour but the psychologist's training covers the whole range of behaviour so that he sees abnormal behaviour as deviation from the normal. Whether he is using tests and specialized assessment techniques or not the psychologist provides a quite different contribution to the understanding of psychiatric patients from the psychiatrist by reason of his totally different training background.

The psychiatrist, having formed a tentative opinion about a patient, frequently needs more information about certain features of the patient's behaviour. He is aware that he requires more reliable, more objective, and less crude techniques for assessment in order to provide a better basis for decisions and therefore more effective prediction.

The psychologist, being somewhat more aware of the need for objectivity and measurement in the description and explanation of behaviour, provides a valuable contribution to decision making through the techniques he has for just such improved assessment. In addition, psychological tests provide an opportunity for sampling a wider variety of behaviour than is possible in an interview. Thus they often extend the range of information about the patient. Indeed the clinical psychologist's assessment of psychiatric patients is best seen as an extension of the examination undertaken by the psychiatrist.

Objectivity

F. H. Allport (1955) has discussed the problem of objectivity in studying perception. His three main criteria of objectivity can be applied in the field of psychiatry. All three criteria are closely related. If they are all satisfied the result is that different observers will be able to agree on the observations they make.

The first criterion is whether observer involvement is reduced to a minimum. Bias can creep in at the level of behavioural description. One aim in using psychological tests is to provide a situation in which the eliciting and recording of behaviour is as independent of the tester as possible. This is arranged by having standard situations in which patients are given standard instructions and by having testers trained to administer tests in a standard way. In addition, rules for scoring are applied consistently as laid down in test manuals. Where these procedures are followed there is less distortion in the description of the patient's behaviour due to the observer. Unfortunately, observer bias can also creep in at the level of interpretation of observations. This is because the experience of each individual observer is biased and any subjective interpretation he may make will be based on this unrepresentative sample of experience. Tests overcome this limitation. They can be administered to representative samples of the population and the data resulting from this standardization can be used to interpret the score of the individual patient. Norms are provided in the test manuals to facilitate interpretation and in this way the observer is eliminated as a source of bias at the level of interpretation.

The second criterion of objectivity is whether the contact between the observer and the observed behaviour is increased as far as possible. Tests are generally devised so that a performance or a response is elicited which can be recorded for subsequent scrutiny. For example in the Porteus Maze Test of foresight and planning the subject is required to show the way out of mazes which are drawn on paper. By using a pencil to do this he leaves behind him a permanent record of his errors and successes. The recording of test responses is intended to improve observer contact or what Allport calls 'denotation'.

The third criterion is whether the operations defining the situation in which the observations are made, are capable of being repeated by others. This is necessary if

S.C.E.R.T., West Bengal
Date 9. 1. 70
Acc. No. 2147

different observers are to agree about what they observe. The use of test manuals describing the procedures required and instructions to be given, as well as the training of testers, all serve to improve observer agreement and repeatability of results.

It has to be acknowledged that much psychological testing is no more objective than the psychiatric examination. Projective tests, for example, cannot be claimed to have any advantage in objectivity, though they may still be useful for other reasons.

Unfortunately the term objective is often applied loosely to psychological tests. In using it authors do not always state clearly whether they are referring to the response elicited from the subject, the scoring applied to the responses, or the inferences drawn from the responses.

A response made in a questionnaire is a self-evaluation based on introspection and is therefore subjective. The response to perceptual tests such as the Archimedes spiral is a verbal report concerning a perceptual experience and is therefore subjective. A Rorschach response is also a verbal report on a perceptual experience and therefore subjective. But a G.S.R. or speed of tapping includes neither self-evaluation nor verbal report and is objective.

Scoring of test responses may involve a subjective judgement on the part of the scorer. This is usually the case with open-ended tests in which the subject is provided with an opportunity of responding in a great variety of ways. Thus the Rorschach is an open-ended test and the scorer has to attempt to classify the subject's responses. A closed test in which the subject is severely restricted in the responses he may make, as in multiple choice tests, requires no judgement by the scorer. A questionnaire that requires 'yes' or 'no' for an answer to each question can be quite objectively scored with no element of subjectivity involved. Wherever there is room for judgement in scoring there is the chance that different scorers will disagree. Where judgement by the scorer is eliminated then scoring will be the same whoever does it.

132.075
GAT

S.C.E.R.T., West B
Library

Conclusions as to what the responses mean may be made on the basis of clinical judgement. The clinician evaluates a Rorschach protocol so that his inferences are subjective. Where inferences are drawn by actuarial means no judgement is involved and therefore no disagreement arises between observers trained to draw inferences actuarially from the scores given by a subject. Such inference is therefore objective (see chapter 3). Both subjective and objective methods of inference may be applied to responses which may be subjective or objective and which have been scored by either subjective or objective means.

Cattell's use of the term objective is rather idiosyncratic. He uses it 'to distinguish from questionnaires those tests in which the individual actually acts . . . and has his performance measured without being aware what traits are being measured' (1965, p. 372). Cattell classifies as objective projective tests such as the Thematic Apperception Test (T.A.T.) and Rorschach as well as laboratory techniques such as motor coordination, body-sway suggestibility, and G.S.R. Questionnaires he classifies as subjective because the subject is evaluating himself.

The same test may be described as objective by one writer and not objective by another. They are not necessarily disagreeing. They are probably referring to different aspects of the test. This variation in usage of the term however does cause some confusion.

It is clear that psychological tests are used in an effort to increase objectivity in assessment. Although perfect objectivity is never achieved, it can be said by reference to the three criteria mentioned above that one procedure is more objective, and therefore a better basis for prediction, than another.

Measurement

The importance of measurement in psychiatry has been emphasized by Mayer-Gross, Slater and Roth (1960).

> In medicine we have to deal with causes of all kinds, and not only those that are both necessary and sufficient. In searching for a cause of some phenomenon, we are really searching for a quantitative relationship . . . We can, if we like, lose our way in a causal network: but we need not do so if we take into account the quantitative aspect of causation. It is at this point that so much psychiatric thought loses cogency and direction. Quantitatively important causes are tangled in a knot of others whose quantitative relation with the effect we are interested in is slight or entirely unknown (p. 6).

It is often claimed that such intangibles as feelings cannot be measured. But where terms such as 'more than' and 'less than' are used in description then some degree or amount of the quality described is implied. People can be ranked in order of the degree to which they possess some quality even though it is not possible to say exactly how much more of the quality one person has than another. This level of measurement is known as ordinal scaling (Stevens, 1951). It is not as exact as interval scaling in which the distance between the numbers is known; as in the centigrade scale the difference between a temperature of 10°C and 20°C is the same as between 80°C and 90°C. By devices such as scale transformations a semblance of an interval scale may be achieved in psychology but this should not persuade anyone that the measurement of behaviour approaches the level of measurement in physics. Psychologists (who know a thing or two about selling) are given to referring to psychological tests as 'measuring instruments' which are 'precise' and 'exact' and even dignify their use by the term 'psycho-metrics'. Such terms often create a misleading, if unintended, impression. Nevertheless, an attempt at measurement on ordinal scales is better than no attempt at all. Indeed a great deal can be achieved with ordinal measurement despite its limitations.

Measurement is an aid to objectivity. Norms based on scores applied to test responses help the psychologist to draw inferences from the individual patient's test performance independently of the psychologist's own background and experience.

There are dangers in applying scores to behaviour. One is that information which is difficult to score or manipulate statistically may be discarded. This may result in a serious loss of data. Another danger is that the test constructor and user may become obsessed with figures so that he loses sight of their psychological meaningfulness. As G. W. Allport (1961) observes:

> Basic methods of quantification are indispensable for determining the reliability and validity of methods. Our only criticism of quantification comes when it is made the primary objective in research and theory rather than a tool . . . Statistics is a useful adjunct but the field of personality should not be a mere playground for the mathematically gifted (p. 459).

Clinical Psychology as Applied Science

One of the principal tasks of clinical psychologists is to extend the use of measuring devices in psychiatry. The clinical psychologist who is in touch with the research literature is in a position to recognize ways in which techniques, hitherto used in research only, may be applied to clinical problems. In this way research developments are made available to psychiatry. And undoubtedly psychology has a great deal to offer psychiatry.

One measure of the value of any science is in its applicability. Although the potential application in psychiatry is great, it is only slowly that the relevance of the concepts, principles and theories of psychology has become appreciated. It is largely due to clinical psychologists that the findings of psychology in such fields as learning, memory, sensation, perception, thinking, motivation, attitudes, interests, social psychology and physiological psychology are being applied at all. The clinical psychologist is first and foremost a psychologist, consequently he will be useful in so far as he is able to relate his knowledge of psychology to the practical problems of psychiatry.

This information traffic does not flow in just one direction. The opportunity of studying the abnormalities of behaviour has benefited psychology. Much of what is known about personality structure and function has been gained from psychiatric patients and many psychological theories have been born or nurtured in the psychiatric clinic.

The clinical psychologist is both applied scientist and pure scientist. The distinction between applied and pure science is difficult to maintain in the older natural sciences. It is impossible to maintain in the fast developing behavioural sciences. The investigation of practical problems frequently produces valuable 'pure' knowledge. Although the knowledge thus gained has practical implications it is as pure as any generated in the laboratory.

Tolerance of Uncertainty

Considering the complexity of the subject matter, the achievements of psychiatry and psychology are a source of pride and encouragement to those working in these fields. Nevertheless the fact has to be faced that both psychiatry as a branch of medicine, and psychology as a branch of science are still at relatively early stages of development. There is much research to be done in these fields. Without the benefit of that research there will be much uncertainty in clinical practice.

The clinical psychologist has to develop the capacity to tolerate uncertainty if he is to continue to work in the clinic. This is not to be taken as a recommendation that he should be satisfied with this uncertain state of affairs. Nevertheless in clinical work it must be acknowledged that much of the time one has to live with uncertainty. Psychologists at times have to make judgements 'on the basis of inadequate data. Time and other economic factors are of considerable importance in clinical work, and decisions of importance may have to be made when hypotheses are untested or partially tested' (Jones, 1960, p. 780). It is

> legitimate to decide to work on the basis of a certain assumption with a patient because the patient's condition urgently demands action and the psychologist or psychiatrist knowingly takes a step into the dark. There are relative degrees of certainty with which one can entertain theories about a patient's condition and what can be done about it ... the applied scientist has often to plunge into guesswork and trial and error ... (Shapiro, 1951, p. 760).

One of the reasons for uncertainty is that the technical aspects of tests (standardization, norms, reliability and validity) are rarely as adequate as might be hoped for. In addition, information about the population to which the tests are being applied (base rates, see chapter 6) is rarely available. Sometimes it is possible to estimate the margin of error associated with a test result but often this is not possible.

Despite these deficiencies the psychologist is often the best equipped person to assess certain aspects of patients' behaviour and he should not avoid making his contribution. As Piercy (1959) points out, if he declines the task it will be undertaken by someone else, using even more inadequate data. The estimates or predictions made by the psychologist are not free from error. Yet they are often the best informed estimates so that the margin of error will be less than is contained in the estimates of others.

The recommendation of a rigorous scientific approach in clinical psychology whilst laudable in its aim often has the unfortunate side effect of producing a reluctance to make judgements and present opinions on the basis of inconclusive data. In medicine decisions are often taken in the face of uncertainty. A doctor learns to take such decisions and to tolerate the uncertainty, whereas the psychologist does not have the same kind of experience. When he enters clinical psychology he is liable to be consumed with doubts and misgivings about the evidence and particularly the validity of the evidence on which many decisions are based. These doubts may be so crippling that he can make no decision at all. Such a person would probably find projective techniques unpalatable and even the limitations of more objective, simpler, better validated tests might generate too much anxiety. These doubts are often particularly severe in the psychologist who has entered the clinical field with a crusading spirit, to free psychiatrists from the dangers of their own subjectivity. He is liable to become disenchanted when he finds that it is often difficult to assess the validity of tests, particularly where construct validity is concerned. Even when correlation coefficients are available they are generally rather low: of the order of 0·20 to 0·60. It is important to point out, as Cronbach (1961) does, that although the validity of a test may not be as great as one might wish, the test will be useful as long as its validity is greater than that of other methods used for the same purpose. Cronbach and Gleser (1965) argue that 'A test of validity 0·20 in one situation may be more beneficial than a test of 0·60 in another. The characteristics of the specific decision determine "What the test is worth".'

Another reaction to this dilemma of having to make judgements on the basis of inadequate data is for the psychologist to minimize the inadequacies. In this case the psychologist denies the uncertainty and acquires a confidence in his judgements which may be quite unwarranted.

Apart from the technical shortcomings of tests there are uncertainties about theories and orientations in both psychology and psychiatry. Desai (1965) points out,

> In the field of personality we have widely divergent approaches, with antitheses between the nomothetic and the idiographic, the atomistic and the gestalt, the molecular and the molar. We have approaches which believe in marked emphases, to say the least, on one or more of the various aspects of personality; for instance, in terms of broad categories, on the mask or the structure aspects of personality, or more specifically, on overt behaviour, traits, hierarchy, integration, dynamism, adjust-

> ment, or uniqueness as the essence of personality. There are approaches to personality which reject, out of hand, any reference to unconscious motives and dynamics while others lay extreme emphasis on unconscious motives and conflicts and minimize the importance of conscious motives. We have beliefs which stress the importance of the independence of contemporaneous motives from their sources of origin. On the other hand, we have emphasis on instinctual urges and early unconscious motives which are believed to continue to influence our apparently most contemporaneous and varied modes of experience and behaviour.
>
> If we bear these facts in mind it becomes obvious that what problems a clinical psychologist poses himself, the way he approaches them, the methods he uses to answer them, the practical inferences he draws from his findings and the sort of predictions he makes, would depend on the clinical psychologist's own frames of reference, his orientations, attitudes and beliefs in relation to general psychology, clinical psychology, psychiatry and more specifically on the views of personality he explicity or implicitly holds (p. 2).

Whilst recognizing the shortcomings and limitations of assessment procedures as well as the theoretical controversies in the field of personality, it is the view of this writer that enough is known concerning assessment for clinical psychologists to make a valuable, independent, and unique contribution to psychiatric problems.

2 Assessment Referrals

Introduction

It is an astonishing fact that there is virtually no discussion in the literature of the kind of referral problems to which psychologists have a contribution to make. One would expect to find in books describing tests used in clinical psychology some mention of what they might be used for. But this is rarely the case.

Since clinical psychologists work mainly in conjunction with psychiatrists it is important that psychiatrists should be informed as to what problems are most profitably submitted to psychological investigation. Obviously if psychiatrists do not ask the right kinds of questions they do not get the answers they need from psychologists. The aim of this chapter is to present briefly the main areas of assessment problems. Before doing this however it will be useful to take a look at the general context within which referrals to the psychologist are made.

Factors affecting referral questions

First, referral problems will vary according to the type of setting in which the psychologist finds himself because of differences in psychiatric decisions required. In outpatient units and admission wards the emphasis is on turnover of numbers, early diagnosis and quick disposal. Since diagnosis in the early stages of mental illness is often uncertain and the effectiveness of treatment is nearly always greatest in the early stages there tend to be a large number of diagnostic problems referred to the psychologists in such units. Where there is less concern about rate of turnover psychiatrists are more prepared to 'wait and see' when diagnosis is uncertain. Most psychiatric conditions will declare themselves one way or the other eventually and a psychologist working in a unit with this sort of approach may get fewer requests for aid in establishing diagnosis and perhaps more for treatment and assessment of treatment outcome.

Second, the orientation of the referent will influence the kinds of problems he refers. If he is keen to initiate psychotherapy he may ask for elucidation of conflicts, psychosexual development and the patient's typical defence mechanisms. If he is interested in neurological aspects of psychiatry he may be on the look-out for cognitive impairment resulting from brain damage so that his referral questions will be different.

Third, the psychologist's own interests will determine the sort of problems he encourages his colleagues to refer. He may be interested in global personality assessment using projective techniques, or he may be more concerned with the study of specific mechanisms and use strictly objective procedures. The sort of problems referred to him will also partly depend on how far he has managed to persuade his colleagues of the value of his particular approach to assessment.

Lastly, referral questions will depend on the assessment techniques available. There are continual developments in this field. Modifications of existing techniques (Kornrich, 1965), novel applications of traditional techniques (Payne and Friedlander, 1962) and creation of new procedures (Foulds, 1962) all have to be kept under review and tried out. When a psychologist is convinced that a research development has clinical application it is for him to demonstrate its utility to his colleagues and persuade them of its value.

Problems associated with referral

There are several difficulties to which both psychiatrist and psychologist should be alerted so that when the psychologist accepts a referral this is the result of a considered decision made in full awareness of the dangers involved.

Junior psychiatrists often have clinical problems which should be discussed with their consultants but which instead get referred to the psychologist. This may happen where channels of communication between consultant and junior are poor or supervision is inadequate.

Where the psychiatrist's case load is very heavy he may unload some of the problems he would otherwise handle himself on to the psychologist. This may not in itself be a bad thing because there is in fact a great deal of overlap of functions. Some kind of division of labour has to follow if each is to work most effectively. The actual division that takes place can be regarded as an administrative problem of staffing but the psychologist should consider carefully the implications of taking on the role of non-medical psychiatrist thrust upon him by such exigencies.

Indiscriminate referrals can be made for the purpose of case conferences. The doctor presenting the case asks for psychological assessment for the sake of completeness and in anticipation of questions which may be raised in the discussion of the case. Such referrals may have educational value and can be used by the psychologist as opportunities for showing his colleagues the sort of problems he wishes to have referred.

Indiscriminate referring of patients to psychologists can occur in units where medical laboratory services are readily available and doctors may get into the habit of referring patients routinely for various laboratory tests. This attitude can be generalized to the psychologist with the danger that the more automatic the habit of referring the more vague is the problem referred.

Inadequate discussion between psychiatrist and psychologist can often produce difficulties. Sometimes a psychologist may reformulate a referral problem or he may unearth a problem during his investigations which he feels is worth further elucidation. For example, a phobic woman is referred for assessment of intellectual level in relation to the demands of her job. The question is whether the stress of her work has to be considered as a factor contributing to her neurotic state. The psychologist to whom the patient is referred is more interested in other aspects of the case and by means of projective tests he discovers sexual difficulties which he identifies as the determining factor in her phobia. This may be something the psychiatrist knows already or something he does not wish to confront the patient with anyway. In such cases discussion with the psychiatrist is essential if the psychologist's contribution is to be clinically relevant. Of course the psychologist does not withhold some observation or finding merely because the psychiatrist did not ask for it. A symptom such as an intermittent Parkinsonian tremor may occur during interview with the psychologist which had not been reported in the case notes. It would be an unwarranted omission if he failed to mention this in his report.

Referrals can be phrased in such a way that the opinion of the referent is plainly stated: for example, 'Would you confirm my impression of defective intelligence?' The psychiatrist in formulating the referral in this way is clearly seeking support for his own opinion rather than an impartial assessment. He runs the danger of encouraging the psychologist (particularly the inexperienced one) to tell him what he wants to know rather than what he should know. This danger becomes important when the psychologist departs from findings derived from the application of properly validated, objective, standardized tests and begins to use clinical judgement and acumen in order to give an opinion rather than a finding. He then runs the risk of his judgement being influenced by the way in which the referral was made. Because there are few tests which are acceptable technically – particularly in the assessment of personality – this risk tends to be a frequent occupational hazard. There is less danger if the psychologist's report clearly states which conclusions are objective findings, which are subjective inferences from the test data and which are opinions less firmly based on evidence.

A further danger the psychologist faces is that he may get swamped with referrals so that he may feel he has to sacrifice any research he may wish to do. However valuable the contribution he makes to clinical problems it is essential that he devote time to reviewing developments in assessment techniques and contributing to those developments. This is necessary if the contribution which psychology has to make to psychiatry is to be fulfilled.

The following discussion indicates some of the ways in which psychological assessment can be useful in work with adult psychiatric patients. No attempt is made to discuss systematically all the referrals which may arise in work with children and in other psychiatric sub-specialties.

Cognitive Deficits

Among the cognitive functions which are frequently impaired in psychiatric conditions are reasoning, learning, memory, mental and motor speed and perceptual-motor coordination. Adequate description of a patient's level of functioning in these areas is often very useful in answering a number of questions which are asked about patients.

Organic or functional?

This is a differential diagnosis frequently in the mind of the psychiatrist and it is a problem on which the psychologist can usually provide valuable evidence. The distinction between depression and dementia in old age; the exclusion of an organic condition in a depressed patient being considered for E.C.T.; the relative severity of organic and functional impairment in a patient with a known organic condition exhibiting some hysterical behaviour; Ganser symptoms requiring identification as manifestations of a schizophrenic, hysterical or organic condition;

the establishment of organic causes in a patient complaining of memory loss or giving evidence of confabulation: these are some examples of such problems.

Suspected organic involvement

Frequently organic involvement of the central nervous system is suspected on the basis of clinical evidence. For instance, change of personality in an elderly person showing loss of social graces and inability to curb aggressive impulses may suggest an organic disease. Unusual word usage may suggest dysphasia. A history of head injury, alcoholism or fits in a patient may arouse suspicion of deterioration. Detailed investigation of such patients with suspected deterioration will be advisable and the results may indicate the need for further physical investigations.

Baseline for serial assessment

Because it is sometimes desired to follow the changes in a patient's mental state over a period of time, repeated assessment may be requested. This is particularly useful in cases of suspected deterioration. Where later testing can be related to earlier performance, changes are more easily detected. The effects of treatment can be followed in this way too. For instance, where a vitamin deficiency is suspected as the cause of cognitive impairment this can be checked if treatment for the deficiency is accompanied by assessment before and after treatment.

Residual impairment after treatment

When symptomatic recovery appears clinically to have gone as far as it will after withdrawal of alcohol, arrest of G.P.I. (general paralysis of the insane), leucotomy, myxoedema or carbon monoxide poisoning for instance, there is usually the question as to whether there is any residual deficit and if so how extensive it is.

Localization

Despite the contradictory findings of localization studies

(Meyer, 1960) psychological tests can often be of use in providing a description of behaviour for the purpose of relating psychological deficits to specific areas of brain damage.

The Use of the Intelligence Quotient

Intelligence is a dimension of personality that has to be estimated in every case a psychiatrist sees if his clinical formulation of the patient's condition is to be adequate. One of the first judgements he will make concerns the patient's level of understanding and expression so that he can adjust his questions and the form of the interview accordingly. During his training the psychiatrist learns to assess intellectual level on the basis of the patient's school record, work record, current general knowledge and vocabulary usage. It is *only* when the psychiatrist cannot do this himself or has doubts about his clinical impression that he should refer a patient to the psychologist *if* it is important to have an accurate indication of the intellectual level of functioning. Such difficulties or doubts commonly arise with the older working class patient whose educational opportunities were poor, in distinguishing within the I.Q. range 75–95 between the categories low average, dullard and borderline defective and with dysphasic patients, deaf mutes or any patients with whom verbal communication is difficult. Suitability for verbal forms of psychotherapy is often partly decided on the basis of intellectual level, so that when there is difficulty in encouraging an inhibited patient to speak during the initial clinical assessment, referral for testing may be indicated.

Educational subnormality is often suspected by a psychiatrist and he may require a more accurate assessment of level of functioning than he can make clinically. Sometimes this may be one of a number of possibilities in his mind. The distinctions between amentia, dementia and simple schizophrenia, for instance, are often not easy to make clinically and testing may narrow down the possibilities.

Court cases are often referred for assessment of intellectual level particularly where a low level is suspected and the question arises of the patient's understanding of right and wrong actions.

There are intelligence tests which psychiatrists can administer themselves. These are mainly tests devised primarily for group testing and are therefore largely self-administering such as Raven's Progressive Matrices and the Mill Hill Vocabulary (Raven, 1958b, 1960). The danger with tests which are easy to administer is that interpretation may be mistakenly thought to be equally easy. Unless a psychiatrist has been adequately trained in the use of the test therefore he is well advised to leave the interpretation to a psychologist.

The psychologist needs some indication of intellectual level in nearly every case. He may need it in order to decide what tests to give. For example some questionnaires are of little value for patients of below average intelligence because of their difficulty in working out the double negatives involved as in the following item: 'I cannot concentrate my mind on one thing.' If you answer 'false' for this you are denying the first negative, which requires an average degree of intelligence to comprehend.

It is important to know the intellectual level in order to interpret results on many tests. A patient's answers to proverbs which are above his intellectual capacity and therefore ones which he could not be expected to get right, would be impossible to evaluate without an I.Q. estimate. Many such tests as the Memory for Designs Test indeed require an I.Q. component, such as vocabulary score, for calculating results.

Classifying and Explaining Abnormalities of Behaviour

In the previous sections on cognitive deficits and assessment of intellectual level the contribution of psychological assessment is obvious, and the problems involved are fairly clearly understood. Development is merely a matter of refining techniques.

In the field of the functional and personality disorders psychological assessment is much more controversial. This is partly the result of the controversies within clinical psychiatry and partly the result of inadequate techniques available in clinical psychology.

Diagnosis

A primary task of the psychiatrist is to classify the patient and put a label or a diagnosis on him. This is often difficult to do with certainty and so the psychologist is often asked to help in the differential diagnosis of patients. One way in which psychologists have attempted to answer referral questions of this sort is by using tests which have been given to groups of patients having the diagnoses in question. Thus Meehl and Dahlstrom (1960) used the M.M.P.I. to differentiate neurotics and psychotics. Payne (1958) has argued that this usage of psychological tests contributes little of value to psychiatric practice. He points out that a diagnosis has a number of implications.

First it implies a description. A diagnosis is used as a kind of shorthand for referring to a cluster of symptoms. Unfortunately one cannot tell from the diagnosis which symptoms are present nor how severe is the condition. A test validated against the diagnosis 'psychotic depression' therefore could not give any more adequate description than is implied in that diagnosis. Payne suggests that the descriptive significance of a diagnostic test score is nil because it adds nothing to the descriptive information of the psychiatrist.

The second implication is for prognosis. Unfortunately few psychological tests are validated against the future course of the illness so that for most tests the prognostic implications are unknown.

The third implication of the diagnostic label is for treatment. Again few tests have been correlated with treatment outcomes so that generally treatment implications will also be unknown.

The fourth implication of a diagnosis is for etiology. A test validated against a diagnostic label can embody nothing more in its scores than is known about the etiology of that condition. In psychiatry little is known of etiology so once more the test implications are severely limited.

Payne concludes,

> The diagnostic test score is thus completely useless. It can be used validly to predict which label the psychiatrist who took part in the validity study would have assigned to a patient . . . it cannot be used validly to predict any of the *consequences* of this label. But the consequences of the label are all that the doctor is really concerned with (1958, p. 27).

Future research may enable tests to predict these consequences but for the present it has to be accepted that diagnostic psychological testing of the kind outlined contributes little to psychiatric practice except in certain restricted areas such as the identification of behaviour associated with brain disorder.

This discussion shows that to ask for a diagnosis is not a very sensible referral for a psychiatrist to make. He should consider the diagnostic implications which interest him. When psychiatrists start asking psychologists for assistance in these areas then psychologists might be stimulated to more appropriate research than trying to predict what labels psychiatrists will put on patients.

Williams (1965) suggests that the best contribution the psychologist can make to the problem of differential diagnosis is to provide the detailed analysis, measurement and description of symptoms which psychological tests permit. The psychologist is on dangerous ground if he makes diagnostic inferences from his test findings. Diagnosis, it is suggested, is the job of the psychiatrist and he can use the psychologist's test findings to put the patient into whatever diagnostic category he sees fit. This recommendation is particularly important where the nosological concepts of psychologists and psychiatrists are at variance. They may agree about the basic data but disagree about their interpretation.

A further objection to diagnostic psychological testing is that overlap of test scores given by patients in different diagnostic categories is generally so great that a case which is difficult for the psychiatrist to classify will also be difficult for the psychologist. This argument is based on the assumption that the psychologist is functioning, as Meehl (1954) has described, as a statistical clerk. Having obtained the patient's score he looks up a frequency table which indicates the probability of the patient falling into one or other of the diagnostic categories being considered (see chapters 3 and 6). Since few tests have been developed to the stage where this is possible, psychologists in practice use the same clinical methods of inference that psychiatrists use. This does not mean that psychologists are operating as substitute psychiatrists. The data evaluated by psychologists are mostly derived from psychological tests and not from the psychiatric interview.

Explaining behaviour

Much behaviour which is difficult to explain is referred to the psychiatrist who may pass the problem on to the psychologist for any light he may throw on the matter. Examples of such behaviour are compulsive gambling, cyclical vomiting in adolescence, indecent exposure, anorexia nervosa, dypsomania, psoriasis and school refusal.

The psychologist may consider the problem from the point of view of abnormal symptomatology or personality assessment, or both. What he seeks is evidence as to whether or not there are psychological factors which will help explain the behaviour and so lead to an indication for treatment.

Jones (1960) gives examples of investigations of individual patients with various forms of disorder including neuroses, behaviour disorders and psychosomatic conditions. The experimental method of setting up hypotheses, and deriving deductions which are then tested under controlled conditions is shown to be applicable to such diverse problems as lack of persistence in a patient with a behaviour disorder, disturbed relationship with the mother in an asthmatic patient and sexual identity in an anorexic patient (see also chapter 3).

Personality assessment

Apart from investigating specific hypotheses of disordered behaviour it is sometimes necessary to arrive at an overall personality assessment, for example in assessing suitability for psychotherapy or in giving a prognosis. For this purpose the psychologist has techniques for investigating (i) the patient's personality strengths and assets as well as his weaknesses and liabilities, (ii) his characteristic defence mechanisms and methods for dealing with stress and conflict, (iii) his self-concept and his attitudes towards himself as well as his attitudes towards other people and his interpersonal relationships, (iv) his needs, fears and aspirations as well as the specific stresses to which he is vulnerable and (v) his interests.

Assessment may be required when behaviour therapy is being considered for a patient. The term behaviour therapy embraces a great number of techniques and regimes aimed at changing abnormal behaviour patterns. If a patient is to be allocated to the most appropriate programme for his individual needs then his case should be properly assessed.

Vocational Guidance

In giving vocational guidance the psychiatrist is frequently in the dilemma of trying to help the patient to make the maximum use of his potential yet at the same time wishing to protect him from environmental pressures which will prove too great for that potential. The psychologist is in a position to assess employment potential through aptitude tests and elucidation of motivation and interests. He can assess whether the patient's personality is compatible with his occupational interests and preferences.

Psychologists' Attitudes to Referral Questions

It is important that the limitations of psychological tests should be freely acknowledged. Unfortunately some users of tests (particularly of projective tests such as the Rorschach) give the impression of a kind of omniscience in psychiatric matters. In addition psychiatrists are occasionally over-impressed by the claims made for tests or by the use of symbols and numbers in test reports. Sometimes the psychologist can give a clear solution to a referral problem. Often, he may only be able to provide information which may not be of immediate value in assisting a decision about a patient but which may be of value at some later review of the case. Often tests are of little value and the psychologist has to plead ignorance. The enthusiasm of the psychologist to demonstrate the value of his tests should not lead him to go beyond their limitations in solving referral problems.

Occasionally a psychologist may feel that his work has been wasted when his findings on a case appear to be ignored by his colleagues. This could occur when communications between psychiatrist and psychologist are poor. The psychologist may then not appreciate the problem the psychiatrist had in mind when referring the case or the psychiatrist may not be sufficiently acquainted with clinical psychology to appreciate the relevance and significance of the psychologist's report. A different situation arises when the psychologist's evidence is discrepant with the psychiatrist's own formulation of a case. This may lead to a rethinking of the case. But if the psychologist's evidence cannot be reconciled with the other evidence available then the psychiatrist is taught to trust his own judgement. A doctor is trained to give credit to his judgement and to consider evidence from different sources in the light of the whole. This is necessary as long as he is also trained to have insight into the common human failing to look for evidence which supports and to ignore that which conflicts with one's hypotheses.

Formulating Problems for Referral

There is a type of referral which is clearly to be discouraged. This is the vague referral such as 'personality assessment' or 'psychological assessment'. Anyone referring in this way has not thought out clearly what kind of information he wants from the psychologist.

Psychiatrists should be encouraged to formulate any questions which they put to psychologists as precisely and specifically as possible. The way in which the psychologist chooses to solve the problem is of course left for his own decision. Where a clear formulation of a problem is made then the psychologist's contribution becomes more relevant and to that extent more valuable. It is probably an advantage to the psychiatrist if he has to formulate as clearly as possible the problem which he is putting to the psychologist because in doing so he should be stimulated to review the case.

In practice the formulation of specific problems is usually the outcome of discussions between psychiatrist and psychologist. What starts out as a general discussion about a typical psychiatric decision such as treatment or disposal gets refined to some problem to which the psychologist can make a distinct contribution. The vague blanket referral is often a plea for help which discussion can translate into a specific problem.

In considering a referral the psychiatrist should ask himself whether psychological assessment will contribute anything to practical decisions about the patient which he will have to make. The only occasions when it is justifiable to refer a patient to the psychologist when it is clear that no practical consequences will follow his assessment are in teaching and research. This means that the psychiatrist should formulate clearly for himself the decision on which the referral has a bearing and he should make sure the psychologist is aware of the relation of the referral questions to that decision.

This will help the psychologist to select the best and most appropriate test procedure for his assessment. A request for an I.Q. may be made as we have seen for a number of reasons. If it is for assessing a patient's suitability for psychotherapy a vocabulary test such as the Mill Hill Test would be used. If the problem was one of vocational guidance another test might be more appropriate. The question might be whether the patient had the intellectual capacity for a university course in which case the A.H.5 would be more appropriate than the Mill Hill Vocabulary.

3 Methods of Assessment

The first step in assessing a patient is to define the problem. This was the subject of the last chapter. The next step is to lay down a scheme for solving the problem. This involves deciding what information is required and how that information is to be gained. The final step is to draw conclusions from this information. What does it signify? Does it provide a solution to the problem?

An examination of the strategies adopted to elicit information and the methods of drawing inferences from that information reveals three methods for investigating patients: the statistical, the clinical and the experimental. In actual practice a psychologist may use any one or any combination of these methods in dealing with a particular assessment problem. In this chapter the main features of the three approaches to assessment are outlined.

Statistical Inference

Description

The main features of this method have been described by Meehl (1954). The aim is to produce a test with standardization such that when administered to a patient, his score can be checked against the tabulated scores of standardization groups and he can be classified accordingly. When research has produced the kind of data for this method to be adequate there is no subjective interpretation of results and no use of clinical acumen in the classification of an individual patient's score. Once the patient has produced a score the classification of that score can be done by anyone trained to read the tables – a person Meehl calls a statistical clerk.

It should be pointed out that it is the mode of drawing inferences from the data which is described as statistical, not the nature of the data which may be objective or subjective, quantitative or qualitative.

An illustration of the actuarial method of inference is provided by the Nufferno Tests of intellectual speed and level (Furneaux, 1956). It is known that schizophrenics are abnormally slow at problem solving. This fact can be used to distinguish normals, neurotics and schizophrenics. For any given level of intelligence as defined by the accuracy score, the relative frequencies of normals, neurotics and schizophrenics achieving various speed scores can be determined. Thus for the accuracy score range 10–12 the ratio of schizophrenics to normals scoring in the speed range 160–164 is 9 : 1. The probability of a patient who scores in these ranges being schizophrenic is therefore 0·9. Similarly it can be shown that the probability of his being neurotic is 0·67.

Once a patient's scores on the accuracy and speed tests are known, anyone trained to use the table could arrive at the probability of the patient belonging to one group or another. In addition all users of the table should come to the same answer, barring clerical errors.

The second illustration involves the use of M.M.P.I. profile signs to distinguish psychotics from neurotics. It was originally hoped by the constructors of the M.M.P.I. that the scales developed from comparing groups of psychiatric patients with normals would help in the diagnosis of patients. This hope was not fulfilled. Instead of, say, the schizophrenics having a single peak on the schizophrenia scale it turned out that they often had several high points. The next step was to look at the profiles and try to distinguish clinical groups on the basis of certain signs.

At first this was done impressionistically. Later workers attempted to make explicit the rules on which their impressions were based. Meehl and Dahlstrom (1960) discuss a set of sixteen rules (presented in Marks and Seeman,

1963) which are arranged in sequence for identifying psychotic and neurotic M.M.P.I. profiles.

> One reads through the list of rules in order until one comes to one which covers the profile under consideration. Application of the rule to this profile then results in one of four outcomes: Profile 'psychotic' (*P*), profile 'neurotic' (*N*), profile indeterminate (*I*), or 'proceed to next relevant rule'. In this manner all profiles are finally classifiable under one of the first three rubrics, *P*, *N*, or *I*. . . . The rules, while complex . . . can, after a little practice, be applied by a clerk in less than a minute's time.

In an extensive cross validation study (N=988) 30% of the profiles were classified as indeterminate with regard to the psychotic-neurotic distinction. Of the rest, 76% were correctly identified. This level of discrimination is superior to the impressionistic profile assessment of individual clinical psychologists.

Once again anyone trained to apply the rules can classify a test record and all users of the rules should come to the same conclusion for each record.

Statistical formulae have been derived (Payne and Jones, 1957; Rosen, 1966) for the purpose of comparing test scores. Psychologists are frequently interested in changes in a patient's test scores over a period of time and also in discrepancies between scores on different tests taken by a patient. It does not matter whether the test itself is a personality questionnaire, an intelligence test, an interest inventory or whatever. As long as the tests provide continuous scores and as long as the required normative statistics are available, such as means, standard deviations and test-retest correlations, these statistical methods can be applied.

Consider the first problem of changes in a patient's test scores over a period of time. There may be reasons for thinking that a patient has deteriorated or improved in some way and tests are called upon to check this. A neurotic patient who has been treated may show a difference in some questionnaire measure of emotional lability given before and after treatment. The question is whether the change is larger than one might expect considering the errors involved in the measurement. We know from test–retest reliability coefficients that psychological tests are not perfectly reliable. If we can calculate the extent of the test error then we can say whether a change in score is within the limits of that error or not.

Rosen (1966) gives an example of a patient whose M.M.P.I. Ego Strength (Es) Scale score on admission was 34 and after four months of psychotherapy was 40. Is this a genuine improvement or could the change in score be explained on the basis of practice effect and test error? Rosen reports the means and standard deviations for a group of patients tested twice with the Es Scale, together with the test–retest correlation. He demonstrates that 19% of patients can be expected to improve their Es scores by six or more points without psychotherapy. One must therefore conclude that the increase in this patient's score lies within the limits of test error and does not indicate significant improvement. Payne and Jones give an example of a schizophrenic patient's changes on the Wechsler Bellevue Test following drug treatment.

The second problem concerning test scores is the comparison of a patient's scores on different tests or subtests. On the Wechsler Adult Intelligence Scale (W.A.I.S.) for example one may find a discrepancy between the Verbal and Performance Scales. Is this due to test error or is it a reliable discrepancy? Field (1960) has provided tables to answer this particular question without one having to do any calculation. Suppose a man aged thirty years has a V.I.Q. of 106 and P.I.Q. of 96 giving a V.–P. discrepancy of ten points. Field's tables show that the probability of obtaining such a discrepancy by chance is less than 0·05. We can conclude therefore that such a discrepancy is reliable and not merely due to test error.

A related problem which frequently arises from test results is whether a discrepancy is abnormal or not. For instance how many people achieve a V.–P. discrepancy of ten points or more? Field's tables once again provide the answer ready made. At least 12% of people in the same age group as the thirty-year-old man have a verbal bias of ten or more points. It is left to the psychologist to decide what psychological significance this conveys. To achieve statistical significance at the usual 0·05 level the V.–P. difference would have to reach seventeen points.

For this patient the conclusion can be drawn that his verbal bias is reliably ascertained although it is not sufficiently unusual to classify it as abnormal.

Evaluation

As noted in the discussion of measurement in chapter 1, there is inevitably a loss of information in reducing behavioural descriptions to scores for the purpose of actuarial processing. This may be valuable if the discarded information is irrelevant but if it is significant and relevant then its loss may result in over-simplification.

Piercy (1959) in discussing the assessment of intellectual impairment has shown that it would not be possible to quantify and build into a formula all the possible variations in behaviour that the clinician observes. The task is too complex and the clinician, it is argued, remains the most essential tool in assessment.

> Now it might *in theory* be possible to devise a battery of tests standardized as to administration, scoring and interpretation, which would accommodate the complexity of test procedure necessary to identify most cases of organic intellectual impairment. It must be stressed, however, that in most cases of mild intellectual impairment the logical structure of any thorough assessment of the evidence is extremely complex, and if formalized would result in a test battery incorporating a formidable decision procedure for administering successive tests. This decision procedure would be of prohibitive complexity and, in addition, there would be a need for an extremely complex system of weightings if the various possible combina-

> tions of response patterns are to be recognized and assessed within the framework of the scoring system. All this might, in theory, be done but only in the same sense that it is possible to formulate a decision procedure for playing the perfect game of chess (p. 493).

A formula or actuarial technique should only be used in conditions which are the same as those under which the technique was developed. If the conditions are different then the technique will give erroneous results. The analysis of W.A.I.S. profiles illustrates this. Field's tables are based on the American standardization data given in the manual. Since no British standardization data are available the use of these tables in Britain is really an act of faith.

The trouble with cut-and-dried tests which give simple scores, easily processed in a mechanical way, is that they have an aura of correctness about them. The danger lies in their facile application in cases where they are not justified. Paradoxically those who accept uncritically the findings from such procedures are often those who regard themselves as rigorous hard-headed scientists.

The major difference in conditions between standardization sample and testing sample is likely to be in the base rates (see chapter 6). If this difference is ignored the test may give grossly misleading results.

Uncontrolled variables (such as the base rates) may change over the period in which a test is used. The test may then become inappropriate even though it was appropriate when first introduced. In order to keep it up to date it should be revalidated and overhauled frequently. This requires a great investment in man-hours which would have to be justified. A judgement would have to be made in each situation as to whether the benefit from the procedure justifies the investment necessary to make it work.

As it happens, the statistical method has not been used in clinical psychology as much as it could be, for precisely the reason that the basic data collection is an arduous business. This means that there are few tests for which adequate data are available for the individual patient's score to be processed actuarially.

There are two major advantages of statistical methods. First, they are more likely to lead to appropriate decisions than the clinical method. Second, they lead to economies since, once the patient's score has been obtained, this can be processed by clerical assistants or even computer. Meehl (1954) says,

> It is my personal hunch . . . that a very considerable fraction of clinical time is being irrationally expended in the attempt to do, by dynamic formulations and staff conferences, selective and prognostic jobs that could be done more efficiently, in a small fraction of the clinical time, and by less skilled and lower paid personnel through the systematic and persistent cultivation of complex (but still clerical) statistical methods. This would free the skilled clinician for therapy and research, for both of which skilled time is so sorely needed (p. vii).

Clinical Inference

Description

When no statistical data are available the clinical method is usually used. This method of analysing data and drawing inferences is the same as that used by physicians and psychiatrists. It involves evaluation by the clinician of the mass of information which is available to him about the patient. He processes this information and arrives at an opinion about the patient's condition. In this procedure of evaluation the psychologist compares the available information with his knowledge of such conditions. This knowledge will be learned from textbooks, from the research literature, from teachers and associates, and from his own experience of similar cases.

The process of evaluation involves continual formulation and reformulation. As each piece of information becomes available the clinician considers its possible implications in relation to all the other information he has received. In diagnosis he considers the different possibilities and attaches some probability estimate to them. If the patient rotates designs in the Bender-Gestalt Test the likelihood of his condition being organic increases.

Hypotheses are tested by searching for information which will confirm or reject them. Gradually by a sequence of such decisions and checks a final judgement is made. In the process some hypotheses will have been considered and rejected.

The whole extended process may be a reasoned argument which is capable of being made explicit for public communication or it may be intuitive with steps that the clinician cannot make explicit.

This method has been described as an art.

> In this approach the clinician behaves like the craftsman in that he attends quickly to what he knows to be significant, his skill develops with experience, and his work involves economy of effort. Like other crafts, it is dependent upon the assimilation of a body of knowledge which is shared by others; but expertness increases with practice until the practitioner is capable of recognizing clinically significant features with such rapidity that his steps of reasoning are not discernible to the uninitiated (Rodger and Mowbray, 1961, p. 167).

At this stage of development the craftsman or clinician may not even be aware of all the features to which he responds. Because of this the clinician's skill has been regarded as lacking the qualities of objectivity and communicability associated with scientific thinking.

The clinical method of inference involves comparing the patient's set of data with the configurations of criterion groups such as schizophrenics, brain damaged, depressed, etc. This is a matching operation with the established configurations acting as blueprints and involves examining test results with respect to their degree of approximation to known syndromes of disability.

The data evaluated by the psychologist are not restricted to test responses. He may draw on the case history, his own interview and general observations of the patient. He may use actuarial evidence described in the previous section where this is available, but generally this will not be sufficient and information from other sources will require to be evaluated as well.

Evaluation

The fact that experienced clinicians frequently disagree in their opinions on the same data is a measure of the subjectivity of clinical inference. One often hears it said that the Rorschach findings on a patient were such-and-such. To be correct one should refer to this particular tester's findings on this patient's Rorschach. Where tests are interpreted by clinical judgement it is as important to know who tested the patient and interpreted the results as to know which test was used.

One reason for disagreement between clinicians is that inferences are heavily influenced by the clinician's theoretical views. It is therefore inevitable that data will be put to different uses by different clinicians. Bias can creep in from several directions: the terms of the initial referral, reliance on memory and the tendency to ignore evidence which is contradictory to one's clinical impression and to select evidence which confirms it.

Another diffculty is that the agreed and substantiated body of knowledge which clinicians acquire through training is very limited. Sometimes what they learn is without scientific support and although it may start off as surmise, with time these speculations achieve the status of accepted fact. This makes it difficult for the clinician to remain with his feet firmly grounded in fact.

As a result of these forces the psychologist using the clinical method of inference faces the danger of going beyond what is justified from a patient's results. More than that he may acquire a confidence in making clinical inferences which, if subjected to a validation study, might be shown to be unwarranted. To check these tendencies clinical judgements should be validated wherever possible. This is necessary especially where the steps in the reasoning process cannot be made explicit. It has been demonstrated that clinicians' judgements can range from accurate and useful to inaccurate and misleading. In other words, when clinical usefulness is checked some clinicians are found to be good and others bad. This exercise in validation of clinical judgement is a salutary one which clinicians should perform to calibrate themselves and to improve their functioning. Unfortunately such a validation study requires following up patients which is often difficult for the busy clinician to arrange.

The use of the clinical method in clinical psychology has been criticized as merely leading to a further psychiatric opinion. The implication is that the psychologist is duplicating the psychiatrist's work. Although this can happen, generally a psychologist is basing his opinion on information which is not available to the psychiatrist. Another criticism of the clinical method is that it tends to encourage the psychologist to accept the psychiatric conceptual framework so that his distinctive viewpoint as a psychologist is abandoned.

Psychiatrists should not be misled into thinking that clinical psychologists use the statistical method when more often than not they use clinical methods of inference. When the psychiatrist first comes into contact with psychological tests he is, as Zangwill (1952) has said, 'perhaps unduly impressed by the prestige of numbers' and he assumes that this is the way psychologists work, objectively and without recourse to the clinician's art. The word 'psychometrics' conveys to him that the ideal state of statistical prediction has, in fact, arrived. When he realizes that it has not and that the psychologist uses clinical experience and judgement, he is just a little disappointed. The danger is two-fold. In the first place he may accept test results in an uncritical manner, and in the second he may come to reject them because they do not do what he expected.

Experimental Method in the Investigation of the Individual Patient

Description

The clinical method in general and particularly in psychology has very low status in the eyes of the advocates of scientific method, despite the fact that the processes are in many respects similar. Shapiro (1951) has suggested that there are considerable benefits to be gained from making explicit the clinician's implicit sequence of hypothesis formulation. He suggests that in so doing the formulation becomes more systematic and potential hypotheses are considered more impartially and carefully. Although the essential characteristics of the two methods are basically the same, the experimental method avoids many of the dangers of the clinical method and opens up new possibilities for investigation and application of psychological knowledge.

The procedure involves first of all a survey of the information available from all sources in the light of the particular problem referred. Explanations may be required for the patient's behaviour and symptoms, for test discrepancies or for test observations. In selecting problems for investigation the psychologist has to keep the referral problem in mind since it is easy to choose problems that are irrelevant to the main issue. This caution does not apply to research-oriented investigations but to those investigations oriented towards the patient's benefit.

Next, all the hypotheses which might be set up to account for the behaviour are formulated. The particular hypotheses that the psychologist considers will depend on his knowledge of psychiatry and psychology as well as his own theoretical orientation, whether it be analytic, neo-Hullian or whatever.

The practical implications of these hypotheses are considered and a decision is made as to which ones to test. This decision will be guided by a consideration of the possible consequences, of each hypothesis being accepted or rejected, for the treatment and disposal of the patient. Will the test of the hypothesis have any practical effect on what is to happen to the patient? If not it is not worth testing.

Often hypotheses will arise from which nothing follows except the facts they seek to explain. These hypotheses will be the first to be discarded. Other hypotheses may have no bearing on treatment and disposal so there is little point in testing them and they can also be discarded. Then certain deductions are made which are tested under controlled conditions. When the results of the testing are known a hypothesis may be rejected, modified or accepted. New problems may be thrown up by the testing and generally a revision of the original hypotheses will be necessary, requiring further testing. In this way a sequence of hypothesis formulation, testing and reformulation will be followed until a satisfactory solution of the referral problem is achieved.

A particular feature of this approach is the emphasis on making controlled observations. The patient may be used as his own control, a single matched control may be chosen, a suitable control group may be used, and finally available norms for tests may be used.

Example

The following case[1] is presented as an illustration of the experimental method in the investigation of the individual patient (Gathercole, 1966a). A man with a long history of obsessional ruminations was judged to be well above average intelligence on the basis of his apparent command of words. He used words like 'metaphorically', 'commiserate' and 'irreparable' which appeared well above the level expected of a person working, as he did, as a hospital porter. One problem put to the psychologist was whether he was underemployed. The second problem was to assess whether there was any evidence of cognitive impairment.

---

[1] Grateful acknowledgement is made to Professor Fish for permission to quote this case.

He was a boxer and it was suggested that his brain might be damaged.

Assessment of his word knowledge was required for an answer to both problems. The first stage in the investigation was to determine whether the psychiatrist's impression of high verbal ability was correct. A standard vocabulary test (from the W.A.I.S.) was administered, on which he scored in the average range. This indicated that his verbal ability was not as high as had been thought. However it might have been that the words which he had used and which impressed the psychiatrist were in fact above average in level of difficulty. In order to compare the level of difficulty of these words with the W.A.I.S. words which he had defined correctly, Thorndike and Lorge's (1944) word count was used. This book gives frequencies of words found in a wide variety of reading matter.

The more difficult words in vocabulary tests are also the more unusual ones. If the words used by the patient included some that were more unusual than those he could define correctly on the W.A.I.S. vocabulary test then the psychiatrist's impression would be confirmed. A list of the impressive words was constructed. Each word had been spontaneously used correctly and later correctly defined by the patient. When these words, as well as the correctly defined W.A.I.S. words, were checked for frequency the psychiatrist's impression was indeed confirmed.

To account for the discrepancy between the frequency of words he used and those in the W.A.I.S. he knew, the hypothesis was put forward that he had acquired a 'specialist' vocabulary. A garage mechanic might acquire a specialist vocabulary of unusual words such as 'sprocket' and 'spline'. These words would not be a representative sample of his knowledge of all possible words of their frequency. It was thought that the patient's preoccupation with words might have resulted in familiarity with a limited number of unusual words.

It follows from the hypothesis that he should have difficulty in defining words randomly selected from the Word Count but matched for frequency with the list of words generated by himself. A second list of words was constructed and presented as a vocabulary test to the patient. The expectation was confirmed and the conclusions were drawn (i) that the W.A.I.S. vocabulary score was a fair indication of his word knowledge and (ii) that his own word usage was misleading.

This simple experiment had important clinical implications. It was important for the assessment of intellectual impairment resulting from brain damage. The base line from which any deficit was to be gauged was not as high as first thought. The experiment was also important for his vocational guidance. It could be stated that his work as a hospital porter was appropriate to his intellectual level.

This example of the experimental method in clinical psychology shows how basic research can develop from the study of a single case. A host of research problems arise in the search for hypotheses and techniques for testing them. The suggestion that the patient's specialist vocabulary resulted from his ruminations was not pursued but would have been enlightening. The discrepancy between frequency of words used in speech and words defined in a standard vocabulary test may have diagnostic value. It might be possible to derive an index of intelligence from frequencies of words used in speech.

Evaluation

This method has been criticized as being so time consuming as to be impracticable (Hamilton, 1964). Whether it is impracticable will partly depend on the need for an early decision for treatment or disposal. It is often said that there are few emergencies in psychiatry, so this criticism may be relevant less often than is supposed. Of course the patient may discharge himself before the investigation is complete, but this is a hazard of any other assessment method. The example given above shows that experimental investigations are not necessarily time consuming.

Although the experimental method has been applied mainly to problems of cognition and learning there is no particular reason why it should be restricted to these areas. Some studies have shown its value in investigations of problems of motivation and temperament, although few have attempted to test dynamic hypotheses.

The method of investigating the hypotheses will depend on the psychologist's ingenuity and creativity as well as his knowledge of experimental method and statistics and suitable techniques. As with any scientific investigation success cannot be guaranteed by using this method.

Choice of Method

The psychologist is faced with the problem of deciding which of these methods to use and when. Since the statistical and clinical methods have received most attention in the literature these will be considered first.

It might be argued that statistical methods are intrinsically superior to clinical methods. In the development of sciences, subjective methods of inference invariably give way to objective methods which are more open to scrutiny and revision. Whilst granting this and acknowledging that we may expect to see a great expansion of statistical methods in clinical psychology nevertheless the decision as to which method should be used in a particular clinical situation must be based on a direct comparison of the efficiency of the two methods (see chapter 6 for discussion of test efficiency). A trial of strength is held in which both methods are applied to a single group of patients. The method that correctly identifies the largest number of patients is declared the winner. In the Meehl and Dahlstrom report, mentioned above, the efficiency of the objective rules for discriminating psychotic from neurotic M.M.P.I. profiles, was compared with the efficiency of twenty-nine Minnesota clinicians. None of the clinicians correctly identified as many patients as did the objective rules.

When Meehl (1954) reviewed studies in which the effectiveness of statistical and clinical methods had been directly compared, he mainly covered examples from vocational selection. Since that time clinical psychologists have been stimulated to compare the two methods in clinical work. It appears that although neither method has done very well, statistical methods are more often found superior to the clinical. However, for many purposes clinical methods are the only ones available.

The experimental method has received much less attention in the literature than it deserves. As psychologists become more familiar with this approach they will become more adept in its use. It will then become more widely known and used. Where time is available and the psychologist can muster sufficient ingenuity to make it worth while it is a method that should be considered. There is a danger that clinical methods may be used where statistical or experimental methods are appropriate merely because of inadequate understanding of these newer approaches. There tends to be a delay between the time that new techniques become available and the time they are introduced into actual practice. Since clinical methods were first on the scene in clinical psychology, those psychologists who grew up with them tend to be reluctant to adopt the newer methods.

In considering which method to use the decision procedure of the psychologist should be as follows:

(1) Where a statistical method is applicable and has been shown to be more efficient than the clinical method, use it.
(2) Where no statistical method is available or those that are available are inferior, consider the experimental method.
(3) If the experimental method cannot be applied use the clinical method.

In actual practice the choice of method tends to be based not on a rational decision procedure like this but on personal prejudice. The psychologist who views himself as a rigorous, hard-headed scientist will prefer statistical methods, not wishing to be associated with crystal-ball-gazing methods of clinical inference. If a psychologist sees himself as a sensitive clinician not wishing to be confined to the role of sausage-machine technician, he will prefer clinical methods. If his basic interest is in research, then he will be attracted by the experimental approach which gives him the opportunity of exercising his research skills and provides leads for opening up fields for research.

4 Conceptualizing the Patient[1]

Introduction

Assessment practice is profoundly influenced by the way in which the behaviour of patients is conceptualized. There are a number of concepts which enter into most discussions in psychiatry either overtly or covertly and it is useful to draw attention to them here. Partly this is a matter of defining terms and usage of words; partly it is a matter of defining universes of discourse. In psychiatry and psychology it is only too easy to use words inconsistently and to shift from one level of discussion to another.

'What is being assessed?' This is the question this chapter attempts to deal with. The aim is not to examine specific tests in order to fathom what they can assess. Nor is it to examine specific characteristics which are assessed by tests. The discussion will be at a more general level concerning basic concepts underlying the use of any test in the assessment of any characteristic. When these concepts are set down side by side it becomes easier to examine the general function of any particular test. Although these are theoretical issues it will become apparent that they have very practical implications for assessment.

The five topics have been arranged as bipolar opposites to contrast them merely for the sale of exposition. It should not be thought that the psychologist is required to make a choice between one extreme and the other. Each is complementary to its fellow, as will become clear.

---

[1] Based on a paper presented to the Liverpool Psychiatric Society, a summary of which appeared in Gathercole (1966c).

Idiographic and Nomothetic Analysis

A basic distinction is to be made between the idiographic and the nomothetic approaches in psychology. Allport (1961) makes this distinction to draw attention to the relative neglect of the idiographic approach which is concerned with the explanation of individual behaviour. The emphasis is on a specific person whose personality organization and trait constellation are unique. In order to explain the individual's behaviour, to predict his response to treatment or to give a prognosis, many different facets of this complex organism, as a unique individual, will have to be considered.

In contrast, the nomothetic approach in psychology indicates an interest in general laws and principles of behaviour applying to all individuals. Since the emphasis here is on groups rather than the individuals that make up the groups there is a concern with norms and group statistics. In this approach generalizations and universal variables or personality dimensions are sought which will reduce the complexities of behaviour as far as possible.

These two approaches are not mutually exclusive; both are necessary and complementary in the study of psychology. This is demonstrated by the interest of Allport in nomothetic research and by the nomotheticists' interest in individuals (Cattell, 1965).

It is not necessary for the clinical psychologist to feel that he has to make a choice between these two approaches. In his daily work he will find that he has to make use of the methods and knowledge of both approaches. Nevertheless, it is useful to distinguish them in order to see in what ways they have influenced developments in clinical psychology.

From the nomothetic point of view the unique individual can be represented as the point of intersection of a number of quantitative variables. His position in this network identifies him uniquely, but this is not to deny that he shares many features of his behaviour with other people. Allport has sharply critized this representation of the uniqueness of individual personality by pointing out that variables interact and cannot be regarded as completely independent. Both Cattell and Eysenck have been quick to acknowledge the need to investigate the interaction of traits (Cattell, 1965; Eysenck, 1966).

The clinical psychologist, therefore, in applying his knowledge of personality dimensions and general principles of psychology to the study of the patient he is concerned with, is constantly searching for ways in which his nomothetic knowledge can be related to that individual patient.

Clinical psychology was for a long time greatly influenced by the nomothetic approach. Tests used with patients tended to be the products of the study of groups. The value of such tests lies in the fact that a patient's test score can be compared with the norms or group data. The use of intelligence tests is an example of the way clinical psychologists studying individual patients have profited from tests devised primarily to study groups. One of the difficulties with such methods is that for purposes of diagnostic classification such tests are generally of little value. This is because there is usually a great amount of overlap of scores between diagnostic groups. Although the mean scores of a group of normals and a group of neurotics on the Neuroticism Scale of the E.P.I. are significantly different, the score of an individual patient will very often be of no help in diagnosis because of the problem of overlap.

There has however been a distinct move towards devising techniques and methods for studying individual patients in their own right. Kelly's repertory grid and Osgood's semantic differential are examples of new techniques for measuring aspects of the individual without reference to group data.

The *Q*-sort technique is another individual method. This has perhaps found its greatest application in the assessment of patient's attitudes towards themselves.

Shapiro's Personal Questionnaire is a method in which each patient generates the items on which he will rate himself on various occasions.

Shapiro has been a consistent proponent of the need to study individual patients in their own right. His advocacy of the application of the methods of experimental science to individual problems is a further development away from interest in groups and towards individuals.

Phenotypes and Genotypes

Psychological assessment is concerned both with describing behaviour and explaining it. The first stage is to describe, and psychological techniques have been devised because they are considered to have certain advantages over the description of behaviour based on the psychiatric interview. Only when the behaviour has been pinned down adequately can one begin to search for explanations of that behaviour. It frequently happens that an observed item of behaviour such as thought blocking can be produced by several causes. The patient may have sufficient insight to stop himself expressing psychotic ideas in case he gets locked up; his attention may suddenly be taken up with something irrelevant; he may go blank so that he has no thoughts to express; he may have too many thoughts to express; he may have too many thoughts and be unable to select from amongst them. The behaviour observed in all these cases may be indistinguishable and yet it is determined by a variety of causes.

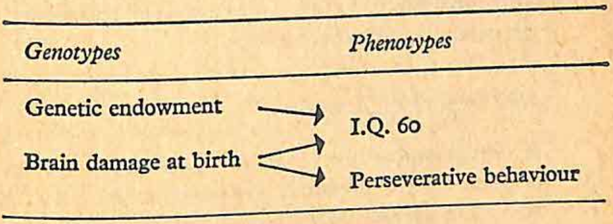
Lewin (quoted in Allport, 1937), borrowing terms from genetics, called the overt behaviour the phenotype and the underlying cause the genotype. The phenotype is analogous to the final common pathway in the nervous system.

A genotype may, of course, give rise to different phenotypes. A bereavement may precipitate mania in one person but depression in another. The experience of failure, which is a genotype, drives one person on to try again until he succeeds but causes another to give up and lose hope. The fire that melts the butter hardens the egg.

The reason why one genotype may have different effects is that genotypes interact. The cause in which we are interested, say the experience of failure, may occur in the presence of a variety of other causes or genotypes which will modify its expression as a phenotype.

It has been noted that psychological tests are useful in describing phenotypes (overt behaviour). The problem here is whether they can help in elucidating genotypes. That is to say, what is their explanatory power?

In the case of mental defect an I.Q. figure will merely indicate the phenotype of mental defect. It will not enable us to distinguish between mental defect due to the normal variation in intelligence and mental defect due to some abnormality such as brain damage resulting from birth injury. In order to do that one has to discover some other phenotypes which are more specifically related to the genotypes being considered. For instance certain types of perseverative response are more likely to be produced by exogenous or pathological mental defectives than endogenous defectives (see diagram).



What one has to do is to decide what the various genotypes or causes may be and then search for some other manifestations which are specific to each genotype and see whether they are present or not.

Another example of the search for genotypes in the field of intelligence is the case in which one attempts to distinguish a person's intellectual capacity or potential from his educational opportunity and cultural background. Intellectual potential cannot be directly measured divorced from environmental influence. But one can estimate the roles of these two influences from a consideration of schooling and family background in conjunction with tests such as the Progessive Matrices and Mill Hill Vocabulary.

Although such estimates making use of test data are undoubtedly more accurate than could be achieved without the data, there is still a margin of error involved.

In the field of temperament, knowledge is very much more limited so that test data are lacking which might help to provide an estimate of the relative influences of constitutional factors such as predisposition to anxiety and of environmental factors such as stress.

As an example consider the case of a spastic girl who easily gets depressed and anxious in the face of frustration and stress. Her emotional control is limited. Is this a direct consequence of brain damage or is it the result of difficulties which her spastic condition has produced?

The examples given above show that psychological tests can be extremely useful in explaining some forms of behaviour, but useless in explaining others. Where they contribute little, this is not necessarily because they are inherently inappropriate. Generally it is because the necessary research has not been done.

Abnormality and Disease

A psychological test is useful in so far as the behaviour elicited can be classified and used to predict something about the person tested. It is therefore important to know what sorts of classes of behaviour are being used in psychiatry. In other words: what does a test response signify?

Frequently psychologists will be attempting to decide whether a test response signifies normality or abnormality, illness or health. It is obviously important to have a clear idea as to what these terms mean.

There are currently held a variety of views on whether psychiatry is concerned with disease entities or dimensions of personality. With some conditions a disease has been established which accounts for the behavioural disorder. G.P.I. presents a wide variety of behavioural symptoms, but these are determined in part by the previous personality and in part by the spirochaete. Many psychiatrists are expecting or hoping for something similar to the spirochaete to turn up to account for schizophrenia.

Traditionally, medically orientated psychiatrists have sought to classify and categorize. If there were no etiological grounds for classification then attempts were made to classify on the basis of the clinical picture. Syndromes or symptom clusters were sought which it was hoped would serve until etiological justification was forthcoming.

A number of writers have been questioning the usefulness of this search for disease entities and seek to delineate abnormalities of behaviour by reference to universal personality dimensions particularly in the field of the so-called functional disorders.

These problems are central to psychiatry and attempts at description and explanation of behaviour will be greatly affected by the view espoused. The kind of techniques devised for assessment will depend on whether a dimensional or a disease entity view is taken.

The history of the M.M.P.I. illustrates the difficulties involved. It was devised '... to provide an estimate of symptomatic syndromes commonly recognized among persons with clinical problems.... The Scales were not expected to measure pure traits nor represent discrete etiological or prognostic entities' (Hathaway and McKinley, 1951, p. 22). The outcome was that items concerned with normally distributed personality traits were mixed in with items concerning physical and nervous disorders. The interpreter of an M.M.P.I. protocol is liable to become confused because he has to try to work in both a dimensional and a disease entity framework at the same time.

How is the term 'abnormal' used in connexion with behaviour which can be classified on the one hand as the result of illness or on the other as deviation from the norm along a dimension? To answer this question it is necessary to look at the three ways in which the term abnormal can be used in psychiatry.

In the first meaning of the term a state which is statistically unusual is regarded as abnormal. Thus being clever is as abnormal as being dull in this usage. In the second meaning some ideal standard is implied such that falling short of this standard is abnormal. These ideals are usually socially determined. Thus being clever is regarded as desirable but being dull is undesirable and therefore only dullness is abnormal in this usage of the term. In the third meaning some disorder of function is identifiable which defines an abnormal state. Thus some defectives are dull by reason of some pathological condition such as phenylketonuria or microcephaly. Had they not been so afflicted their intellectual potential might well have been higher. The brain dysfunction therefore produced psychological dysfunction. These pathological defectives have also been called exogenous defectives to distinguish them from the endogenous or sub-cultural defectives in whom the subnormality of intelligence is due merely to the usual processes of distribution of the genes which govern intellectual potential in the healthy population. No malfunction is present in the endogenous defectives so they are not abnormal in the third sense.

These three different meanings of the term abnormal are frequently confused. It is possible of course for a person to be abnormal in all three senses. A pathological defective for instance is (a) statistically unusual, (b) falls short of the social ideal and (c) exhibits disorders of intellectual function.

Further confusion arises when the terms disease, illness, sickness, pathological and morbid are equated with any but the last of these meanings of abnormality. Attempts to clarify these concepts have been made by Jaspers (1913), Schneider (1958), Szasz (1960), and Eysenck (1960), amongst others. These writers distinguish on one hand those disorders of behaviour which are concomitant with organic changes, such as in delirium and, on the other, those disorders which cannot be regarded as having a morbid organic cause.

Conditions of the first type involving lesions of the nervous system are regarded by these authors as illnesses and classified as neurological. Conditions of the second type are not regarded as illnesses. Maladjustment or emotional disorder may be better terms to use.

Endogenous or sub-cultural mental defectives may not be able to function in a fluid industrialized society, but they are not regarded as being ill, despite the fact that they come under the aegis of psychiatry. In the same way psychopathic and neurotic people are not ill.

According to Schneider, abnormal personalities are abnormal in the first sense defined above, that is they are statistically unusual. Among this group of abnormal personalities is a subgroup of people called psychopathic personalities[1] which includes two further subgroups: those who suffer (the neurotics) and those who cause others to suffer (the sociopaths). Neither are to be considered as ill or diseased, although their level of adjustment is poor.

The reason why these psychopathic personalities come into contact with psychiatry is that morbid states are known to produce identical patterns of behaviour. Such morbid states may have to be excluded by medical investigations. Thyrotoxicosis often presents itself as anxiety neurosis, porphyria can present itself as hysteria, and neurasthenia can be mimicked by various physical conditions. These conditions are not actually neuroses and have to be classified as diseases.

It can be argued that a predisposition to develop neurotic behaviour may be pathologically determined. In the same way that some people are educationally subnormal by virtue of some pathological state whereas others are so by virtue of normal genetic deviation, the constitutional basis of neurotic predisposition may be distinguished. Some people will become neurotic by virtue of normal processes of genetic distribution of the properties governing neurotic predisposition. Others will become neurotic by virtue of some pathological change which will increase their predisposition to become neurotic. It is not being denied that experience plays a role in the acquisition of neurotic behaviour. The intention here is merely to examine the relevance of the term illness to neurosis. Neurotic behaviour will be the result of an interaction between environmental influences, especially in childhood, and constitutional factors such as the ease with which anxiety or aggression are aroused.

---

[1] Note that Schneider's use of the term psychopathic personality differs from English usage in which the term is generally reserved for the antisocial misfits (Mayer-Gross, Slater and Roth, 1960).

For the purpose of illustration consider the arousal of anxiety. Presumably there is a dimension along which people can be ranged from high anxiety arousal in response to stress to low anxiety arousal in response to stress. This will be constitutionally determined originally and those with high anxiety response to stress will require a lower degree of stress to produce neurotic behaviour.

The constitutional basis of this dimension may be genetically (endogenously) or pathologically (exogenously) determined. This is the same kind of distinction as that made in the classification of mental defectives. There is, then, an interaction between the degree of stress the individual is subjected to and his constitutional make-up.

Dimension of Constitutional Predisposition
Insensitive to stress ⟷ Sensitive to stress
α endogenous *β*
A exogenous *B*

Dimension of environmental stress
Small stress ⟷ Great stress
a *b*

Neurosis may result from:

(1) Predisposition a interacting with stress b
(2) „ A „ „ „ b
(3) „ β „ „ „ a (and any stress greater than a)
(4) „ B „ „ „ a (and any stress greater than a)

In the terms of Schneider and Szasz only (2) and (4) are illnesses resulting from pathological causes.

Eysenck and Eysenck (1964) make the claim that the phenotype – neurotic behaviour – is assessed by the E.P.I. *N* Scale. The genotype – constitutional predisposition, or what the Eysencks refer to as the inherited degree of lability of the autonomic nervous system – is measurable by laboratory techniques such as dark adaptation, motor dexterity, conditioning, G.S.R. and so forth. This claim has not so far achieved wide acceptance and no attempt has been made to distinguish between the various types of constitutional predisposition mentioned above by means of such tests.

Extravagant claims have been made by Cattell, Shagass and Eysenck, and others that such laboratory techniques are of demonstrated value in clinical assessment. Such optimism appears to be stronger in the laboratory than in the clinic, and it seems safe to say that these remain primarily research and not clinical tools at the present time.

Foulds (1961) distinguishes between personality traits on the one hand and symptoms and signs on the other.

> A personality trait is taken to be a relatively consistent and enduring position, in respect to a personality variable (which is universal).
>
> Symptoms and signs of mental illness are distinguishable from traits in that they are neither universal nor necessarily enduring. Indeed they indicate a break in the normal continuity of behaviour (p. 269).

Foulds demonstrates that in the field of neurosis it is useful to distinguish hysterical and obsessional personality traits from hysterical and obsessional symptoms of neurosis. Although the symptoms tend to occur with the traits of the same name, cases occur in which obsessional symptoms are associated with hysterical traits and hysterical symptoms with obsessional traits.

Eysenck (1960) regards neurotic behaviour as merely one extreme of a dimension of neuroticism. Likewise psychosis is considered as lying at one extreme of a psychoticism dimension. These dimensions are regarded as normal personality traits and appear to do away altogether with the disease entity notion in the functional psychiatric states. Adams (1964), in discussing the use of the M.M.P.I., says 'the clinical scales of the M.M.P.I. were initially validated against Kraepelinian disease entity diagnostic criteria. The results of the M.M.P.I. factor analytic studies imply that the Kraepelinian labels presently employed in psychiatric diagnosis are non-existent verbal abstractions' (p. 196).

Clearly there are a number of different views on abnormality and disease in psychiatry and clinical psychology. The psychologist has to try to decide whether he is assessing dimensions of personality or disease entities. Confusion frequently arises because these frames of reference are not distinguished. Users of the M.M.P.I. and Rorschach are particularly exposed to this risk.

Structure and Function

One task that nomothetically inclined researchers such as Cattell have set themselves is to identify the units of personality. They seek the basic traits or dimensions of personality which will permit each person to be described as briefly and as precisely as possible. By means of statistical techniques it has been demonstrated that many aspects of behaviour, attitudes and interests, as well as behavioural acts, are correlated. Factor analysis sorts out those qualities which tend to be associated.

To the factor analyst, 'personality structure' refers to a person's location on the dimensions produced by factor analysis. Cattell (1965) uses sixteen personality factors at present (he has a few more up his sleeve), some of which are combined to produce two second order factors. Eysenck (1960) has three major dimensions. Foulds (1965), eschewing the empiricist approach of factor analysis, has selected on rational grounds one dimension for conceptualizing the relationships between different forms of maladjustment. This dimension consists of increasing degree of failure to maintain mutual personal relationships. It extends from normality through personality disorder, personal illness (neurosis) and psychosis to non-integrated psychosis. In addition he uses a personality trait – hysteroid: obsessoid – as well as several clinical scales.

There are numerous tests on the market purporting to provide dimensional descriptions of individuals (e.g. 16 P.F.). It is worth noting that these tests generally do not provide a direct measure of a dimension. Rather they are enumerations of qualities which contribute to the operational definition of a dimension. Thus a person who acknowledges in a neuroticism questionnaire that he blushes easily, is moody and his feelings are easily hurt is regarded as more neurotic than a person who does not have these experiences. This is rather like identifying a tall person as one who acknowledges that he usually finds himself looking down at people he talks to and has difficulty in finding space for his knees in cinema seats. There is no direct measure of neuroticism as there is for height, so reliance must be placed on such indirect measures as adding up different qualities which subjects admit they possess.

If it is accepted that a person's behaviour can be usefully described in terms of his location on a number of dimensions, the question arises as to what is to be done with this description in the case of the individual patient. The problem of combining and integrating the information provided by a dimensional description for the purposes of making decisions about patients is one for which only partially satisfactory solutions have been devised. The clinical method may be used, in which case the profile is subjectively evaluated by the clinician. Statistical methods have also been developed for the purpose of integrating a patient's scores on several dimensions. Cattell, who is foremost in this field, has described several methods (Cattell and Eber, 1957). The principle of profile matching is to compare the patient's profile with those of criterion groups such as anxiety hysterics, schizophrenics and so on. Two techniques of profile matching are available. In the first an estimate of the goodness of fit of the given individual's total profile to any criterion profile is made. This is done by calculating a profile similarity coefficient – r_p. It is claimed that such coefficients can be used to assign a patient to a typical syndrome category. In the second technique for profile matching, a 'qualification grid' is available for each diagnostic category. This grid gives weights for each factor score. A patient's factor weights corresponding to his factor scores can be added up to give a total weighted score which is used to indicate the likelihood of the patient's belonging to any particular diagnostic category.

Apart from profile matching, a technique for 'criterion estimation by the specification equation' is described. This is like a multiple regression equation.

These statistical techniques have been introduced relatively recently and it is too early to say how useful in clinical work they may become. One difficulty appears to be the shortage of suitable reference group profiles and data.

In addition to the factor-analytic approach to dimension definition there is a developmental or genetic approach. Behaviour can be considered as varying along an evolutionary scale from the primitive, undifferentiated, global behaviour of the infant to the individuated, differentiated behaviour of the mature person. In childhood growth, there is evolution, progression or individuation. The child moves through a sequence of developmental stages. At each stage new skills and abilities become differentiated but also integrated in the upward movement toward maturity.

Psychiatry is often concerned with people in whom this process has been reversed so that instead of evolution there is dissolution, instead of progression there is regression as in a dementing process. In addition psychiatry is concerned with people who have not developed to full maturity, whose behaviour has become fixated at a lower level. Homosexual attachment may be regarded as a normal phase of sexual development but a person may become fixated at this level and fail to progress to heterosexual relationships.

It is often wrongly thought that genetic psychologists claim the pattern of personality in regression to be the same as the pattern in the developing child. The young child, the senile dement and the deteriorated schizophrenic may all produce perseverative behaviour. In this respect all three are alike and it is instructive to compare them in this particular. But there will be many differences in other respects.

The writings of Jackson (1938), Piaget (1929), Freud (1959), Werner (1948) and Goldstein (1939) are all directed to elucidating the stages of development in different aspects of personality and to the mechanisms of evolution and dissolution. Significantly, these writers have used mainly idiographic methods of investigation.

There are then a number of ways in which the psychiatric patient can be conceptualized: firstly in terms of empirically derived factors, secondly in terms of rationally defined dimensions and thirdly in terms of an evolutionary scale of development.

This concern with dimensions and factors has gradually given way to a concern with mechanisms and processes. Indeed, if disordered behaviour is to be regarded as a disturbance of functioning, then the mechanisms involved would seem worthy of study. This need was recognized by Rapaport (1945) who deplored the lack of theoretical context of psychological tests and commented, 'the realm of psychological functions tapped by testing procedures . . . has barely been touched by systematic investigations' (vol. 1, p. 4). Twenty years later psychologists are still crying this message in the wilderness (Williams, 1965; Tolor and Schulberg, 1963). Williams points out that clinical psychologists have been too involved in devising tests and subsequently applying them to a wide variety of subjects. Frequently this has been done merely to see if a test will expose differences between psychiatric groups and normals. This kind of work has generally arisen from the demands of psychiatrists for diagnostic tests which will aid classification of patients. Often it is quite devoid of any theoretical justification.

The extensive use of factor analysis has led to the belief that there are 'pure' psychological functions and that 'pure' or homogeneous tests of these functions can be devised.

The functions arrived at by factor analysis are abstractions and can be viewed in different ways. Intellectual functioning, for instance, has been shown by Furneaux (1960) to be usefully regarded as being made up of three components – speed, persistence and level. There are individual differences in the level of difficulty or ceiling to which people can rise in solving problems when given unlimited time. In addition, some people are slower in problem solving than others even though they may achieve the same level of difficulty if given unlimited time. Persistence varies too. The person who gives up easily does not give himself the same chance of achieving a higher score as the tenacious person.

The Block Design sub-test of the W.A.I.S. (Wechsler, 1955) is regarded as a reasonably pure measure of intelligence (correlation coefficients are around 0·70 with full scale score). There are, in fact, a number of processes involved in constructing block designs. First there is the perceptual aspect – reception of the design. Then there is the visual analysis and lastly the executive or motor aspect. In addition, memory, set and attention are involved. Breakdown could occur at any point in the chain of processes. The phenotype (error in construction) gives no indication as to the location of the genotype.

> Psychological tests can no longer be regarded as tests of specific restricted functions. They should rather be conceived as inducing sample performances, or specimens of behaviour, in which some aspect of the person, say his problem-solving ability, has designedly been given prominence, but to which many other forces in his personality contribute. We should recognize that we can never arrange a situation in which just one important variable is tested. Our problem-solving test will perhaps be also a test of frustration tolerance, a test of control over anxiety, a test of level of aspiration. . . . What we say about our subject's problem solving thus becomes a statement about his observed problem solving under a specific set of circumstances, external and internal. We have not measured his capacity; we have merely taken a specimen of his use of his capacity under particular conditions (White, 1952, p. 10).

Quantitative and Qualitative Changes

The difference between behaviour which is regarded as normal and that which is regarded as psychiatrically abnormal may be either quantitative or qualitative. Foulds makes the disconcerting comment that 'psychologists can produce continua like rabbits out of a hat by an artificial sleight of mensuration' (1965, p. 211). By what criteria does one assess whether behaviour is continuous or discrete?

Many psychiatric symptoms are apparently so alien to the normal person's experience that they have always been regarded as qualitatively different. Hallucinations, delusions and conversion symptoms are examples. However, examination shows that most symptoms can be produced in normal people given the right conditions: isolated people such as lone voyagers hallucinate, dreams and hypnagogic phenomena are not far removed from hallucinations, drugs such as the psychotomimetics can provoke psychotic experiences, stress such as in war can induce symptoms of neurosis and psychosis. Comparison of subjective experiences such as imagery (McKellar, 1957) suggests that the difference between normal and abnormal functioning is one of degree.

In longitudinal observation of patients the transition from one condition to another may be traced, so that what appear to be two qualitatively different states may be only extremes of a continuum. Loss of insight is often sudden but it may be gradual, and where it is gradual the change can be seen as one of degree. Where these transitions are rapid it appears as if a qualitative change has taken place.

Mechanisms used by normals, such as repression or dissociation, may be used excessively by say hysterics and this can be seen as a quantitative difference.

These considerations suggest that in many cases what at first sight appear to be qualitative differences may in fact be quantitative.

It is possible that both quantitative and qualitative changes may take place together. This is not uncommon in nature: for instance the physical changes from ice to water at 0°C and water to steam at 100°C. The quantitative change is in temperature and the qualitative change is in the physical state.

If a genetic or developmental view is adopted then it becomes apparent that with growth and individuation new qualities emerge. Raven for instance has spoken of the development of eductive reasoning or the ability to reason by analogy as a salmon leap in the river of life. A new quality of thinking has emerged. Clearly intelligence involves qualitative as well as quantitative differences. The emphasis on intelligence as a quantitative dimension tends to obscure the qualitative differences.

It was pointed out above that many psychological dimensions are measured by adding up different qualities as for instance in questionnaires. In the discussion of phenotypes and genotypes it was pointed out that a bit of behaviour, such as that involved in getting a score on an intelligence test, might be produced in different ways. Thus a different combination of qualities could produce the overt behaviour we are interested in. For example, on the dimension of intelligence three different people achieve a Wechsler I.Q. of 65.

The first is a defective whose scores are:

Verbal Scale I.Q.	65
Performance Scale I.Q.	65
Full Scale I.Q.	65

The second is brain damaged whose scores are:

Verbal Scale I.Q.	80
Performance Scale I.Q.	50
Full Scale I.Q.	65

The third is an illiterate delinquent with these scores:

Verbal Scale I.Q.	50
Performance Scale I.Q.	80
Full Scale I.Q.	65

These people are each located on the dimension of intelligence at the same point – I.Q. 65. But this figure is made up in different ways and conceals these important differences.

Emotional lability, as measured by questionnaires such as the Manifest Anxiety Scale, the E.P.I. *N* Scale and the 16 P.F. second – order anxiety factor, appears to draw on several components. Again a score on the general factor may conceal important differences in the pattern of scores on the specific factors.

Foulds (1961) made an important contribution to the discussion of quantitative and qualitative differences in abnormal behaviour when he distinguished traits and attitudes on the one hand, which emphasize the continuities of behaviour from symptoms and signs on the other, which emphasize disruption of continuity. In making this distinction he is attempting to separate universal personality traits which apply to the total population of normal and abnormal, from those features of maladaptation which are not universal. Although we are more or less honest (personality trait) we do not all have functional paralyses (symptom).

As Foulds (1961) points out

> the distinction is not always simple, we do all wash our hands. The obsessional neurotic may wash his hands more often. This is a quantitative distinction only. The normal person, even if he is of obsessoid personality, feels free to decide whether or not he will wash his hands: the obsessional neurotic feels no such freedom and would be distressed by this compulsion itself or by interference with it. This is a qualitative distinction, which implies either a symptom or a subjective sign (p. 265).

The problem of conceptualizing in terms of dimensions crops up in the assessment of the behavioural effects of brain damage. Psychologists are often asked whether there is evidence of a patient being brain-damaged or not. Many tests are based on a comparison of a group of brain-damaged patients and a group of normals. Scores are derived and the likelihood that a subject falls into one group or the other can be found by comparing his score with the group data. This produces a dimension of 'likelihood of brain damage'. Such a dimension receives some theoretical support from Lashley's law of Mass Action which relates the amount of brain tissue removed to the degree of behavioural disturbance.

In conditions such as senile dementia there is a relatively generalized atrophy of the brain which is related to degree of disorder of behaviour. The inference of mild, moderate and severe brain disorder from mild, moderate and severe deficit on psychological tests in patients with senile dementia is probably justifiable.

This use of dimensions is probably not very helpful. There is evidence to show that a person can undergo extensive brain operations such as leucotomy without appreciable effect on test results. Thus it is the location of the lesion which is often of paramount importance. It follows from this that the psychologist should be concerned with the quality of the performance on tests given to patients in the assessment of the behavioural concomitants of brain damage.

One of the reasons why the quantitative notion of brain damage has been so popular is that tests which give a single score are much easier to use than those which require the painstaking observation and analysis of qualitative aspects of test performance. The exception to this rule is the Rorschach Inkblot technique which can be used in the assessment of brain damage. It requires quite complex analysis yet only provides an indication as to whether the patient is or is not likely to have suffered brain damage. It is of little help in defining the deficits produced by the brain damage.

The need for quantitative and qualitative studies to complement each other is seen in much psychological testing. In their detailed qualitative examination of patients with brain disorder both Goldstein (1939) and Luria (1966) neglected to provide the quantitative normative data for the tests they devised and used. Those psychologists who have invested heavily in developing quantitative tests, such as Payne and Friedlander (1962) with tests of thought disorder, have neglected the detailed qualitative analysis of their patients' responses.

The question of continuity arises in the assessment of psychotic and neurotic depression. Is there a continuum of increasing severity on which these conditions are best thought of as lying, with psychotic depression being more severe than neurotic depression? Or are they best thought of as discrete conditions? Another continuity – discontinuity problem is concerned with paranoid delusions. Many people consider patients with paranoid delusions as related by virtue of their position on a dimension of thinking disorder. At one extreme will be those with delusions but no thought disorder – the paraphrenics or paranoid psychotics. Along the dimension will be found every gradation from mild through moderate to severe disorder of thinking processes. Others consider paranoid psychosis and paraphrenia as quite separate categories from paranoid schizophrenia.

It is probably true to say that every psychiatric condition can be thought of as continuous in some respects and discontinuous in others with every other condition as well as with normality. In other words there are similarities and differences. If one school of thought concentrates on the similarities and another concentrates on the differences they can maintain a controversy indefinitely. What has to be decided is what are the important aspects to be attended to and what are the trivial aspects which can be ignored. Importance will be gauged according to usefulness in treatment and prevention of maladjustment. If dimensions are regarded as convenient fictions we want to know which ones are most convenient for what purposes.

Conclusions

An attempt has been made in this chapter to draw together some of the concepts which underlie the use of psychological tests. In doing this it is hoped that individual tests may be seen in proper perspective.

Setting out these topics in this way also enables one to see what lies at the root of many variations in assessment practice and of many controversies in this field. It also helps one to see what directions future assessment practice will take.

5 Selecting Tests

The Need for Selection of Tests

There are a vast number of tests, techniques and procedures at the psychologist's disposal. In addition to those published commercially there are those described in journals, as well as those developed locally and unpublished. There are many established tools, many which have fallen into disuse, and many which contend to be the popular tools of the future. Some tests are very widely used in practice, but this is no guarantee of their value. It would be unfortunate if choice of tests was affected by their popularity, for there are more rational guides for choosing tests than popularity. In this chapter some of the reasons are discussed why clinical psychologists choose to use some tests rather than others. Given the same assessment problem, two psychologists might use entirely different tests. The aim here is to make the newcomer aware that he is making a choice when he decides to use a particular test. The choice should then be a considered one.

Throughout this book the terms test, technique and clinical instrument have been used interchangeably. The Standing Committee on Test Standards of the British Psychological Society (1965) has suggested certain distinctions, as follows:

(1) A test is any standardized device from which an objective quantitative score is derived.

(2) The word 'test' has been casually used to refer to three different types of device for which an agreed standard usage should be adopted as follows:

- (i) *Tests of attainment*, for example in arithmetic, spelling, grammar, shorthand.
- (ii) *Psychological tests* of cognitive abilities, special aptitudes,

interests, or orectic functions. They include (a) group tests, normally employing pencil and paper, (b) tests designed to be administered individually, which may employ apparatus.

(iii) *Clinical instruments* or quasi-test devices, which permit of a diagnostic assessment of individual performance, but which do not yield an objective score.

The American Psychological Association (1966) in its *Standards for Educational and Psychological Tests and Manuals* suggests that devices for identifying interests and personality traits through self-report should not have the word 'test' in their titles. Rather they should be entitled 'inventories', 'questionnaires', or 'checklists'. No objection is raised to the use of the word 'test' in reference to these devices in textual material for simplicity of language.

The Use of Test Batteries

At the present time a psychologist selects the tests he gives to a patient according to the particular assessment problem posed. He might ask one patient to fill out a questionnaire and do some aptitude tests and another to interpret proverbs and sort objects.

In the 1940s and 1950s – early days in clinical psychology – the use of standard batteries of tests for all problems was common (Rapaport, 1945). Each patient was systematically given say the Wechsler-Bellevue, a Story Recall Test, a Sorting Test, the Rorschach, a word association test and the Thematic Apperception Test (Schafer, 1948). This was justified by drawing attention to the inadequacies of tests. If one test was misleading another would provide a check. Cattell (1965) has drawn attention to this problem in his discussion of instrumentation factors. It appears that certain aspects of behaviour are elicited only by certain tests. For instance several of Cattell's personality factors are only identified by questionnaires and as yet have no counterpart in other types of test. A standard test battery was also regarded as necessary to tap different aspects and levels of the patient's personality. A single test might not be representative of the disturbed level or area. Nowadays many tests, such as Cattell's Objective-Analytic battery, are themselves made up of a number of tests but this is rather different from the use of batteries described above. The use of test batteries may be justified when a psychologist is testing blind. That is, when he is not acquainted with the patient's history and current difficulties. In such a situation the psychologist is expected to track down the diagnosis without assistance from any source other than his tests.

Where the psychologist is asked a question or formulates a problem in the light of all that is known about the patient then there seems no justification for the routine administration of a test battery with the exception of certain situations. In clinical work it may happen that certain tests are very commonly required as, for instance, achievement tests in child psychiatric work. For research, for screening purposes and possibly in the training of clinical psychologists the use of a standard battery of tests may be helpful. In clinical work however there is the danger of using them unthinkingly without consideration of the questions they are being used to answer. Indeed they may be used to answer questions where they are irrelevant or of no practical value. For example when a patient with known memory deficit is referred for assessment of the extent of this deficit it would not be to the point to give the Rorschach. This test might throw light on the patient's condition but would be irrelevant to the assessment of his memory.

The danger of routine application of a test battery only exists in the clinical approach. There is a tendency for the clinician to rely on a few tried and trusted old tests, such as the W.A.I.S., Rorschach, Bender Gestalt and T.A.T. Although many other tests are available the clinician has made a decision to invest his time in these few, and he may get to the stage where he uses them so frequently that his use approaches the battery method.

A clinical psychologist has to make the decision as to whether he is going to invest time and effort being trained and gaining experience with a few tests such as the Rorschach and W.A.I.S. so that he can develop as an expert in the use of the clinical method or whether he is going to widen his experience so as to make a passing acquaintance with many tests. An argument in favour of the former is that only constant usage of a test will make the examiner aware of many pitfalls. Psychiatrists are often happier to rely on the judgement of a psychologist who has a lot of experience with a few tests than on the judgement of a psychologist who knows a lot of tests but has little personal experience in depth of any single test.

The value of knowing a lot of tests is that those which have specific uses can be applied when appropriate problems arise. In testing for malingering it is useful to know of previous approaches to the problem (Jones, 1960). The danger of the second approach is that one becomes a jack of all trades but master of none to the detriment of one's clinical contribution.

Amongst psychologists using the clinical method there is a tendency to reduce the number of tests used, in contrast with those using the statistical and experimental approaches where there is a distinct proliferation of tests, especially of the objective sort.

This can be seen as a case of the bandwidth-fidelity dilemma discussed by Cronbach and Gleser (1965). Does the tester attempt to measure many dimensions crudely or one dimension accurately? An example of a wide bandwidth, low fidelity test is the Rorschach, which is claimed to measure many dimensions including intelligence, emotional control, reality contact, and maturity. Conclusions about any one of these should only be regarded as rough estimates. The Manifest Anxiety Scale is an example of a test having narrow bandwidth but high fidelity. It is specifically designed to measure only one dimension so that the range (bandwidth) of information provided by the M.A.S. is much more limited than that provided by the Rorschach but its accuracy (fidelity) is much greater.

Factors Influencing Choice of Tests

Naturally where available those tests which have achieved acceptable standards of reliability, validity, objectivity and standardization will be given first preference (see chapters in Anastasi, 1961; Cronbach, 1961; Sundberg and Tyler, 1963).

It is worthwhile repeating here the caution that a test which has achieved adequate validity in its development phase may only be used in clinical work with patients of the type used in the standardization. A common clinical use of tests is in cases where the diagnosis is in doubt. Ideally a test used for such a purpose should be standardized on such cases. Follow-up would establish the diagnosis to be used for separating patients into groups of those with and those without the dysfunction being assessed. This procedure was used by Walton and Black (1959) in validating the M.N.W.L.T. Often test developers do not include this phase in the standardizing of their tests. Frequently a test is administered to groups of subjects where the diagnosis is clear. Usually discrimination is good and the test becomes adopted enthusiastically for use in a clinical situation for which appropriate data are not available. The moral is an old one: tests should be used with the population on which they have been standardized.

It must be acknowledged that there are very few tests which are completely acceptable technically, and for clinical use imperfect tools have to be tolerated very often. In assessing the significance of a patient's results the clinician has to take these imperfections into account.

Evaluations of tests are available (Anastasi, 1961; Buros, 1965; Nunnally, 1959) which give guidance on the usefulness of individual tests.

The American Psychological Association (1966) has published a set of *Standards for Educational and Psychological Tests and Manuals* by which test users can judge tests for themselves. If these Standards are applied systematically a test's imperfections are quickly highlighted.

Selection of tests will depend on the type of problem referred. There are some problems, such as the question of organic impairment, for which a variety of tests have been devised. There are others for which tests are available having direct relevance, although not necessarily devised for that purpose. Jones (1960) describes the use of 'level of aspiration' techniques with a man having psychosomatic symptoms and a history of failure at two universities. Situations may arise for which no tests are available when it would be necessary to construct suitable tests *ad hoc*. This would occur more frequently in using the experimental method. The decision as to what tests should be given in a particular case will generally be taken in the light of what is already known about the patient. As each test is given, the situation is reviewed and in the light of the knowledge provided by the tests already given the decision as to which test to use next, if any, will be taken. For instance a man who is known to have cerebral arteriosclerosis is referred for assessment of evidence for cognitive deterioration associated with brain disorder. He scores average on most of the W.A.I.S. subtests and the memory tests given. However on the Similarities subtest he scores significantly lower. This fact now becomes the focus of further enquiry and tests of verbal reasoning are given to establish the importance of the low Similarities subtest score. Had he scored average on that subtest then testing might have been brought to a halt. The actual sequence of tests given to any particular patient will therefore depend not only on the referral problem but also on his performance on each successive test.

Test selection will also depend on the method chosen to tackle the problem posed. If a technique is available with adequate actuarial data for the statistical method of inference to be applied then this should be used. Frequently this will not be the case and one of the other methods will be used. If tests are available which require clinical evaluation and the psychologist has the experience to make this worthwhile he may use the clinical method. Often neither statistical nor clinical methods are applicable and the experimental method may be used if the psychologist feels competent to make it worthwhile.

Choice of tests will be affected by the psychologist's orientation and this will often be derived from the orientation of the department in which he was trained. There appears to be a polarization in preferences between the simple and quantitative as contrasted with the complex and qualitative. In the area of testing for brain damage the work of Kendrick (1965) and Goldstein and Scheerer (1941) illustrates this contrast. Kendrick's data are scores on simple tasks. For example in a Digit Copying Test the patient is given 120 seconds to copy digits. The score is the average time taken to copy one digit. Goldstein and Scheerer are interested in the psychological meaningfulness of the qualitative records of behaviour which they observe. To achieve this patients are given tests which are often open-ended and which provide them with the opportunity of giving infinitely varying responses. For instance in the Sorting Test Goldstein and Scheerer observe a patient's sortings of objects. They record not only the finished sortings but also the development of those sortings and all the behaviour occurring as well. The mass of data collected in this way is in sharp contrast to the single score used by Kendrick.

Choice of tests will also depend on the kind of evidence which is being sought. Sometimes the evidence sought is to be acquired by using a well standardized test, which will provide scores that can be compared with normative data. In this case where an individual's score is to be compared with the standardization norms, it is imperative that the administration should follow the instructions laid down in the manual. Any departure from these instructions means that a direct comparison cannot be made. Sometimes it is necessary to depart from the standard procedure of administration. Examples of such situations are given in chapter 7. When these situations arise the psychologist may be willing to forego the advantages of comparing his patient's data with the norms available in order to acquire other information about the patient which does not require such comparison with norms. This is a dilemma that frequently occurs. Should he sacrifice what may be valuable clinical material in order to get a valid score or should he sacrifice the score in order to gain the clinical material? Some schools of thought avoid this dilemma by insisting on achieving a valid score, even when its information value is trivial. Presumably it is felt that a valid trivial finding is better than a clinical finding, the validity of which is unknown.

The psychologist who modifies a test should of course be aware that he is adapting the test to his own use. The danger arises when he tries to modify test instructions and at the same time interpret the resulting data in the light of the published norms. He should also be aware that his interpretation of data obtained from his modification of the test will require to be validated.

Practical considerations of time will affect test selection in any individual case. For example, in a single outpatient assessment session there will be little time to develop the experimental approach and the statistical method may not be applicable. In selecting tests for clinical evaluation the aim is to try to elicit the information that is most useful in assessing the case. For example, if it is not possible to administer the whole of the W.A.I.S. the decision as to which subtests to use will be determined by an estimate of the relative contribution of the different subtests to the assessment problem posed.

There are a number of tests with overlapping areas of usefulness. In such cases of duplication the customary processes of evolution should eliminate those tests with lower reliability, validity, and poorer standardization. For example, the Bernreuter Personality Inventory has been superseded by tests such as the 16 P.F. However, many tests have special applications and although they overlap with other tests it may be necessary to retain them for these functions. The Coloured Matrices (Raven, 1958a) can be used with a bed patient who is able to do little more than point, whereas the constructional and verbal tasks of the W.A.I.S. would be inapplicable.

Many psychologists hold the view that tests should not be used which interfere with the development of the relationship between patient and psychologist. It is possible for a psychologist to test a patient with hardly any face to face meetings at all. Having formulated a problem from the case record and discussions with the psychiatrist he may select tests which the patient can do on his own, such as questionnaires and inventories, Progressive Matrices or Mill Hill Vocabulary. The T.A.T. often proves more useful when the patient is started off and then allowed to continue on his own, writing out his stories himself. Whilst these methods may be economical of the psychologist's time there is the danger that he may omit to interview the patient altogether.

Testing may influence the patient's attitude toward the psychologist so that for instance a therapeutic relationship may be started off on the wrong track. The test situation structures the relationship in ways that may be undesirable for therapy. For instance the patient is put in the position of having things demanded of him. He is required to submit himself and perform as required by the tester who decides what shall be done. This is in contrast to a psychotherapeutic relationship in which the therapist wishes to present himself in a non-authoritarian role. Once the patient has acquired a set for testing it may be difficult for him to acquire a new set. A similar problem arises when a psychiatrist embarks on psychotherapy with a patient he has examined physically. The relationship of the patient to the therapist is influenced by the physical examination. Where the therapist is active and directs the patient, as is generally the case in the various procedures associated with 'behaviour therapy', then this problem does not arise. It only becomes important when the therapist plays a fairly passive role. One way round this difficulty is to have the patient tested by one person but treated by another.

Test data can be taken as the point of departure in psychotherapy. For instance a patient may be asked to provide his own interpretations of the T.A.T. stories he has told. If this is compared with the tester's analysis, the emerging discussion leads on to therapy.

An extremely important consideration in deciding on whether to use a test is one which paradoxically has nothing to do with the test. It has to do with the population on which the test is to be used. This is the subject of the next chapter.

6 Base Rates and the Usefulness of Tests

The base rate is the actual rate of occurrence of a condition in a particular setting. For example E. Slater (1965) followed up for an average of nine years sixty-one patients receiving the diagnosis of 'hysteria' and in whom no organic condition had been identified as the basis of the hysteria. Twenty-eight (46%) of the sixty-one patients were subsequently found to have organic conditions which would account for the symptoms of hysteria. The base rate of organic conditions in patients diagnosed as 'hysteria' in that hospital at that time was therefore 46%. The probability of a patient being organic would be 0·46. Meehl and Rosen (1955) use the term 'antecedent probability' in their discussion of base rates and the use of psychological tests. Insurance companies determining the frequency of occurrence of events such as death, accidents and so forth, build their data into expectancy tables. These give the base rates or frequencies for these events and are used to calculate premiums.

The interpretation of scores on many psychological tests depends on the base rates prevailing in the clinical situation in which the test is used. If a test is interpreted in ignorance of base rates then use of the test may actually cause more incorrect decisions to be made about patients than would occur without the test.

In order to illustrate the vital importance of finding out base rates wherever tests are used in clinical practice for diagnostic purposes the following hypothetical example is given.

Example

Suppose a test is devised for differentiating the brain-damaged from the non-brain-damaged. In a validation study fifty brain-damaged (B.D.) patients and fifty non-brain-damaged patients (N.B.D.) are given the test. The base rates are therefore 50% B.D. and 50% N.B.D. A cut-off score is found which gives the following results:

Table 1

		Actual condition		
		B.D.	N.B.D.	
Test classification	B.D.	45 (90% of 50)	5 (10% of 50)	pB.D. = 0·9 $\left(\frac{45}{50}\right)$
	N.B.D.	5 (10 % of 50)	45 (90% of 50)	pN.B.D. = 0·9 $\left(\frac{45}{50}\right)$
		50	50	$N = 100$

1. 90% of all patients are correctly identified by the test (90/100).
2. 90% of the B.D. patients are correctly identified by the test (45/50).
3. 90% of the N.B.D. patients are correctly identified by the test (45/50).
4. Of the cases scoring above the cut-off score 90% are in fact B.D.
5. From 4 it follows that the probability that a patient who scores above this cut-off score and whom the test calls B.D. is in fact B.D. is 0·9.
6. Of the cases scoring below the cut-off score 90% are in fact N.B.D.
7. From 6 it follows that the probability that a patient who scores below the cut-off score and whom the test calls N.B.D. is in fact N.B.D. is 0·9.

Suppose that the test is now used with the same cutting score, in a clinical setting where only one patient in ten referred to the psychologist for assessment of brain damage is in fact brain-damaged. The base rates are now 10% B.D. and 90% N.B.D. This might happen in a neurosis unit where the situation could be represented as follows:

Table 2

		Actual condition		
		B.D.	N.B.D.	
Test classification	B.D.	9 (90% of 10)	9 (10% of 90)	pB.D. = 0·5 $\left(\frac{9}{18}\right)$
	N.B.D.	1 (10% of 10)	81 (90% of 90)	pN.B.D. = 0·99 $\left(\frac{81}{82}\right)$
		10	90	$N = 100$

If the psychologist using the test in this situation had not considered the change of base rates from the validation study to the clinical situation, he would be under the impression that the probability of a test positive (cases called B.D. by the test) being valid is still 0·9 whereas in fact it has dropped to 0·5. He would also be unaware that the probability of a test negative (cases called N.B.D. by the test) being valid had risen to 0·99.

In this clinical situation a score below the cutting score could be interpreted with confidence whereas the interpretation of a score above the cutting score is far from clear.

Suppose the test is used in another clinical situation such as a neurology clinic, where nine out of every ten cases are in fact brain damaged. The base rates here are 90% B.D. and 10% N.B.D. With the same cutting score the following table shows the effects of the new base rates on test results.

Table 3

		Actual condition		
		B.D.	N.B.D.	
Test classification	B.D.	81 (90% of 90)	1 (10% of 10)	pB.D. = 0·99 $\left(\frac{81}{82}\right)$
	N.B.D.	9 (10% of 90)	9 (90% of 10)	pN.B.D. = 0·5 $\left(\frac{9}{18}\right)$
		90	10	$N = 100$

The probability of a test positive being valid has now risen to 0·99 whereas the probability of a test negative being valid has dropped to 0·5. Here the meaning of a test positive is clear but the meaning of a test negative is far from clear.

In these two examples it is assumed that the psychologist is testing every patient whereas generally he will see only a selection of patients. The relevant base-rates now are not those applying to the whole clinic. The psychologist will only be interested in the referral base rates which may well be different from the clinic's base rates. Indeed one could go so far as to say that each referring doctor will have his own base rates.

Efficiency of Tests

The vital importance of considering base rates in assessing the usefulness of psychological tests has been shown by Meehl and Rosen to whom reference should be made for a more detailed exposition of the problem. They show how the efficiency of tests can be evaluated in several ways.

Where a dichotic classification is required such as brain-damaged (positive), or not brain-damaged (negative), the data can be arranged in a table in this way.

Table 4

		Criterion classification		
		Positive	Negative	
Test Classification	Positive	a	b	$a + b$
	Negative	c	d	$c + d$
		$a + c$	$b + d$	$a + b + c + d$

1. Efficiency in detecting valid positives is given by $\frac{a}{a + b}$. This indicates the proportion of test positives which are valid. That is to say, the proportion of all those called positive by the test who are in fact positive.

2. Efficiency in detecting valid negatives is given by $\frac{d}{c + d}$. This gives the proportion of negatives which are valid. That is to say the proportion of all those called negative by the test who are in fact negative.

3. Overall efficiency is given by $\frac{a + d}{a + b + c + d}$. This gives the proportion of the total number of cases which are correctly identified by the test.

The example has shown how efficiencies can change when base rates change. In order to maintain a particular level of efficiency it will be necessary to consider changing cutting scores. In the neurosis unit a higher cutting score would be necessary if the efficiency in detecting valid positives is to be kept at 0·9, when base rates drop from 50% B.D. to 10% B.D. Unfortunately this may mean that so few cases score above the new cutting score that in practice it becomes uneconomical to use the test.

To illustrate further the importance of calculating efficiencies of tests the problem of classifying subnormality is presented (Gathercole, 1966b). It is often said that a psychiatric diagnosis of subnormality is not to be identified with intellectual subnormality, yet many people think there is a close correspondence between the two.

A W.A.I.S. I.Q. of 70 has been widely regarded as the cutting point for intellectual subnormality. The example below shows that probably three-quarters of those with I.Q.s below 70 are able to live independent lives without being regarded as psychiatrically subnormal.

The calculations are based on three figures:

(1) an estimated 1% of the total population is psychiatrically subnormal.

(2) 2·2% of the total population score below I.Q. 70.

(3) an estimated 60% of psychiatric subnormals score below I.Q. 70.

This data can be arranged in the usual way as shown in table 5.

Table 5

		Diagnosis		
		Subnormal	*Not subnormal*	
W.A.I.S. I.Q.	69 and below	60% of 1 = 0·6	2·2 − 0·6 = 1·6	2·2
	70 and above	40% of 1 = 0·4	97·8 − 0·4 = 97·4	97·8
		1	99	100

The efficiency in detecting valid positives $\left(\frac{a}{a+b}\right) = \frac{0{\cdot}6}{2{\cdot}2}$ $= 0{\cdot}27$. This means that 27% of test positives are valid and consequently 73% are invalid. In other words three-quarters of the people with I.Q.s in the subnormal range are not psychiatrically subnormal.

Where the base rates are known they can be used to predict the behaviour of individual patients. For instance if eight out of ten psychiatric patients are depressed, one could say that the next patient to be seen will be depressed and one would be correct eight times out of ten. If a test is used to determine this state and the test has an overall efficiency of 0·7 then more mistakes will be made by using the test than by using the base rates. One criterion therefore, of the usefulness of a test is whether it is more or less successful than the base rates. If the test is no more successful than the base rates then the test should not be used. In the example given in the previous section the overall efficiency of the test in both the neurosis unit and the neurology clinic is no greater than the efficiency of the base rates themselves (see tables 2 and 3). If all cases are diagnosed as belonging to the majority group the diagnosis will be correct in 90% of cases. If overall efficiency is the criterion of the test's usefulness, therefore, the test would be dropped since it adds nothing in efficiency of detection.

For any given cutting score, the range of base rates, over which a test is more efficient than the base rates, can be determined. To illustrate this point the graph shown in figure 1 has been prepared from data on the Memory for Designs Test (Graham and Kendall, 1960). If the base rate for brain disorder is known then the minimum cutting score which makes the test more efficient than the base rates is easily found.

Thus if the base rates are 20% B.D. and 80% N.B.D. then the minimum cutting score which can be used is three. This gives 18·6% misclassifications. Using the base rates alone, one would call every patient non-brain-disordered and be wrong 20% of the time. Using a cutting score less than 3 produces more than 20% misclassifications. For example a cutting score of two gives 21·6 misclassifications.

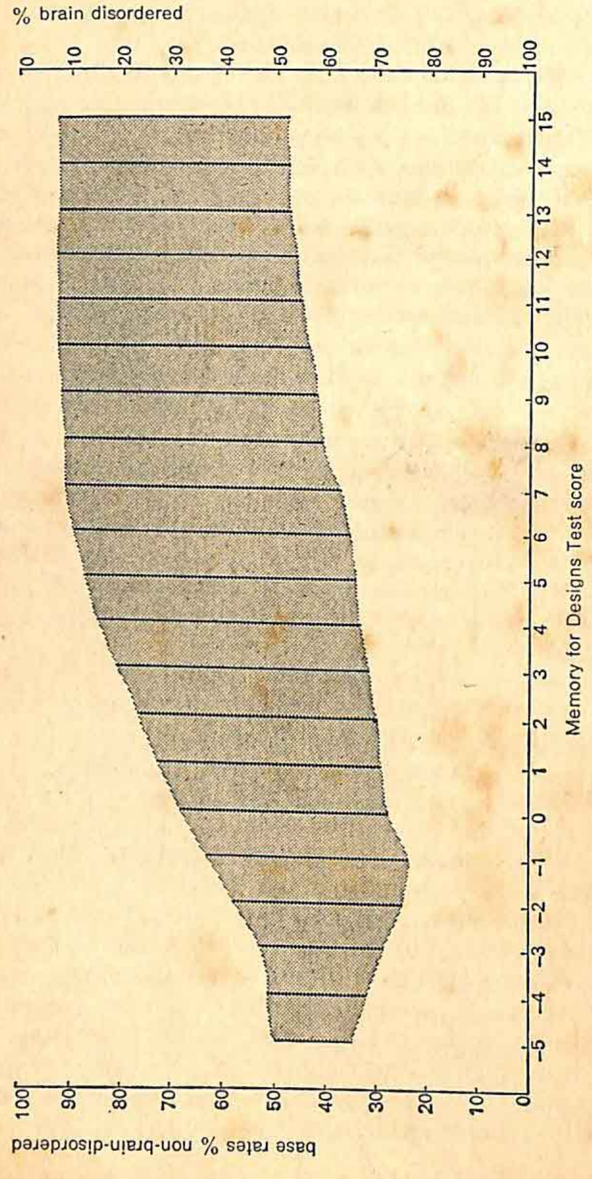


Figure 1 Graph showing the base rates for which the Memory for Designs Test is more efficient than the base rates for different cutting scores.

The graph shows that the overall efficiency of the test is such that the test should not be used if the base rate of non-brain-disordered is less than 25% or greater than 94%. This does not rule out the use of the test in these regions for either detecting valid positives or valid negatives. All it means is that for detecting both valid positives and valid negatives the test is worse than the base rate. Clearly if all the patients are called non-brain-disordered, all the brain-disordered are misclassified. The misclassification of brain-disordered by the test will be considerably less than this.

Test Misclassifications

There is always a margin of error in using psychological tests so that some people will be misdiagnosed. Cases identified by the test as positive which are in fact negative are called false positives. Cases identified as negative by the test which are in fact positives are called false negatives. These misclassifications are called 'misses'. Correctly identified cases are called 'hits'.

In terms of table 4 given above:

a cases are true positives, valid, hits.
b „ „ false positives, misses.
c „ „ false negatives, misses.
d „ „ true negatives, valid, hits.

The decision as to what the cutting score for a test should be in any particular setting will depend on the values attached to the errors made by the test. False negatives may prove so costly that it is worthwhile to attempt to reduce these at the expense of increasing false positives. Thus in testing a patient suspected of an intracranial tumour, a false positive result is regarded as less harmful than a false negative result. If he is pronounced ill when he is well he receives further observation and possibly further physical investigations but eventually his well-being becomes apparent and he is discharged. If he is pronounced well when he is in fact ill he may be discharged without further investigation and attention so that the consequences in this kind of case may be more serious for a false negative than for a false positive.

It is apparent that values enter into these decisions on setting cutting scores. The test user has to weigh the loss from accepting a true negative as positive against the loss of rejecting a true positive.

Problems in Detecting Base Rates

Where the information required to establish the base rates for any condition is available in hospital files, the task of compiling the necessary data is straightforward, if tedious.

Unfortunately the data are often not readily available. The trouble is that there is no general base rate for the whole population which could be used by every tester. Each clinical situation has its own base rates. The example given above (tables 2 and 3) shows that wide differences can occur in base rates between a neurosis unit and a neurology clinic. This means that base rates have to be established locally. Indeed each psychologist should really establish his own personal base rates.

A further problem is encountered in getting follow-up information on patients who have been given psychological tests. Tests are often used in cases where the diagnosis is not clear. In order to find out the base rate of a particular condition, cases accepted for testing would have to be followed up until a diagnosis had become established. This may be as long as two or more years and patients may well have left the hospital by that time. Follow-up information therefore is very often not readily and routinely available and generally is only acquired by setting up a special investigation.

Another problem arises when cases for which a test is used occur infrequently. Where the question of identifying the presence of brain damage only occurs three or four times a year it is obviously going to take a long time before an adequate series can be built up.

Yet another problem is that base rates are often changing over time. Thus a particular doctor who frequently refers cases to the psychologist for say tests of thinking disorder may leave a unit with the effect that the rate of such referrals may drop suddenly. Administrative changes may take place so that patients of a certain type are channelled into another unit. Sometimes a psychologist may be well aware that the base rates are changing. At other times the changes may be unknown to him. In either case it would require continuous checking if changes are to be observed properly. This means that a single study to establish the base rates in a particular clinic could not be relied upon at a later date if conditions have changed.

These points show that for many clinical situations it is not possible to turn the theoretical developments on base rates to practical advantage. If it is possible to determine base rates in a particular situation then they should be determined if tests are to be interpreted properly. The difficulties mentioned above show why knowledge of base rates is generally rather sparse.

In the absence of base rates in a particular clinical setting it may be possible to arrive at an estimate. The literature may give a guide. For instance in the cases diagnosed as hysteria and followed up by E. Slater (1965) the proportion found ultimately to have been brain damaged at the time of the diagnosis of hysteria was 46%. If a psychologist was working in a similar setting he could use this figure as a rough guide in estimating the base rate.

Failing this an estimate may be based on the opinion of the psychiatrist at the time he makes the referral. If he feels he is making a shot in the dark he is more likely to miss than to hit, so the base rate for such a condition would be estimated at less than say 30%. If he has a strong suspicion about a case then the estimate for the base rate would be put at greater than say 70% on the assumption that the psychiatrist is more likely to be right than wrong.

This of course would be a rather hazardous procedure but it should be remembered that without some estimate of base rates a diagnostic test should not be used at all.

Meehl and Rosen (1955) give a formula for determining what the base rates would have to be for a positive diagnostic assertion from the test to be more likely true than false.

$$\frac{P}{Q} > \frac{p_2}{p_1}$$

Where p_1 is the percentage of actual positives called postive by the test

p_2 is the percentage of actual negatives called positive by the test

P is the base rate (percentage of total) of the actual positives

Q is the base rate (percentage of total) of the actual negatives (100-P).

The following example illustrating the use of this formula is taken from Meehl and Rosen (1955).

Piotrowski has presented a set of ten Rorschach signs of brain damage. He suggested that if five or more of the ten signs are present in a patient's Rorschach record then he is probably brain damaged. A cross-validation study by Dorken and Kral found 63% of organic and 30% of a group of non-organic psychiatric patients had Rorschach records with five or more of Piotrowski's signs.

Thus

$$\frac{p_2}{p_1} = \frac{0{\cdot}30}{0{\cdot}63} = 0{\cdot}48$$

and for the conclusion that brain damage is present to be correct more than half the time $\frac{P}{Q}$ must be greater than 0·48. From this P is found to be 0·33. That is, the percentage of actual brain-damaged patients in the population concerned has to be greater than 33% for the efficiency of the test in detecting valid positives to be at least 0·5. If the rate of brain damage is lower than this then test positives will be more often wrong than right.

Attaching Probabilities to Test Scores

Because there is always a margin of error in drawing inferences from test scores it is important to know what is the probability of a particular test indication being correct. This will show how fallible the test's diagnosis is in each individual case.

If a cutting score is used in a test the probability of any single score being correct is given as that for the particular half of the range in which it happens to be (that is, above or below the cutting score). Thus the probability associated with a score just below the cutting score is just the same as the probability associated with the lowest score of all. For clinical purposes therefore the use of cutting scores is often inappropriate. It is usually far more useful to discover the probabilities attaching to each score. This has to be done for a whole range of base rates.

The following graph show this for the Memory for Designs Test (Graham and Kendall, 1960). It was constructed from the cumulative percentages for both control and brain-disordered groups available in the manual for this test. From these percentages it is possible to calculate the percentages of each group making each score. Thus there are 0·5% controls with difference scores of ten or eleven whereas there are 11·1% brain-disordered patients scoring in this score interval. The probability of a score being valid positive can then be determined for any given base rate. Thus for base rates of 10% brain-disordered and 90% controls.

$$p = \frac{11{\cdot}1\% \text{ of } 10}{(11{\cdot}1\% \text{ of } 10) + (0{\cdot}5\% \text{ of } 90)}$$

$$= 0{\cdot}71$$

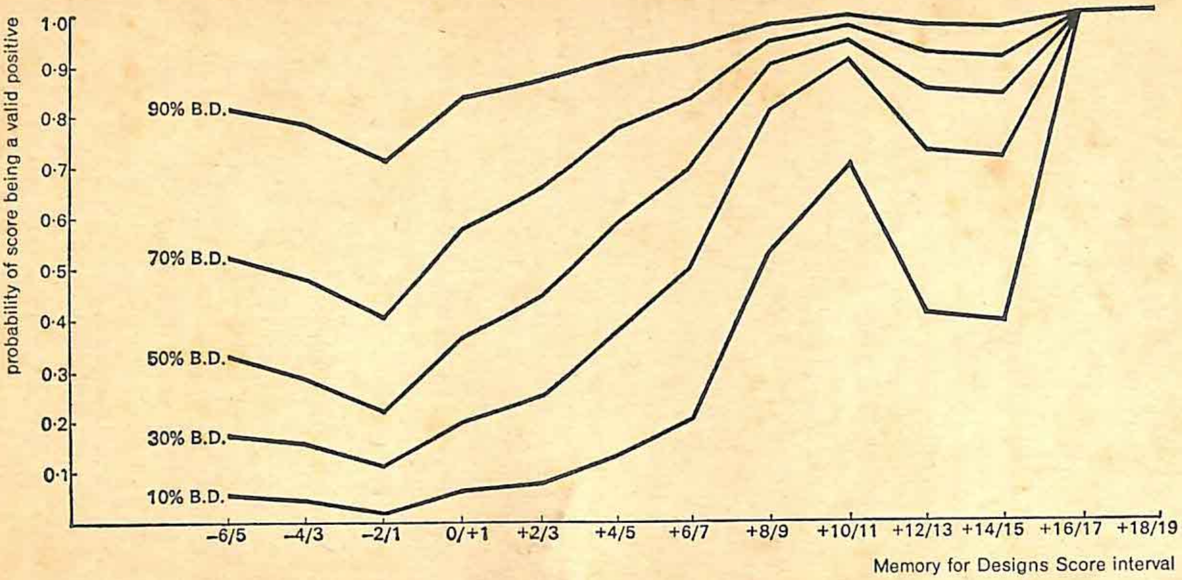


Figure 2 Graph derived from Memory for Designs Test data for control and brain-disordered groups showing probability of score at different intervals giving a valid positive for various base rates.

The statement can then be made that in a situation where only 10% of patients referred for assessment of impairment consistent with brain damage are in fact brain-damaged, a patient scoring ten or eleven has a probability of 0·7 of being brain-damaged. This means that seven out of ten patients scoring ten or eleven would in fact be brain-damaged.

This graph turns out to be very instructive. Taking the base rate of 10% brain-damaged it is seen that a patient has to score at least sixteen before one becomes fairly certain of his being brain-damaged. All the scores from fifteen down to seven, which are all above the official cutting score, give probabilities of brain disorder ranging from about 0·2 to 0·7. Obviously it would be best to call scores indeterminate in this region of probabilities. Graham and Kendall, recognizing that users of the test may wish to calculate probabilities and cutting scores for themselves, provide cumulative percentages for this purpose. It is hoped that in future all test constructors will provide such data for this purpose.

If probability curves are based on a small standardization sample then cross-validation is likely to produce rather different curves because of sampling errors. In order that the curves should be stable on cross-validation, large numbers of subjects must be tested. Even though Graham and Kendall used 243 brain-disordered and 535 control subjects, there were fluctuations in the curves plotted for each individual M.F.D. score. To smooth out these bumps a score interval of two has been used, yet an unexpected dip occurs between scores 10/11 and 16/17. One would expect that as the score rises the probability of brain disorder would rise also. In fact as the score rises from ten to fifteen the probability of brain disorder falls. This dip may be due to inadequate sampling, in which case larger numbers are required to give a more accurate indication of where the curves lie. It may, however, be a genuine dip and if so it illustrates the need for plotting probabilities of actual scores rather than relying on cutting scores. If cutting scores were relied on this dip would not have come to light.

What level of probability should be taken as indicating brain damage? Some suggest 0·95 in line with statistical usage. Thus one would be wrong once in twenty cases and this would be taken as too infrequent to be a chance occurrence. When the consequences of the decision are considered, however, it is clear that 0·95 is much too rigorous a level. Although the number of false positives may be small the number of false negatives (the more serious error) will be large. Even with 0·5 a psychiatrist may feel justified in calling for further investigations because he does not wish to make the error of calling a patient well when he is in fact ill. If psychiatrists accepted the use of probabilities attached to test scores this would be more useful to them than the sort of statements psychologists tend to make in their reports such as: 'The patient scored eight on the M.F.D. Such a score is most frequently found in patients who have suffered brain disorder.'

7 Assessment Technique

Introduction

Despite the guidance provided in test manuals on administration procedure and the instructions to be given to patients, the psychologist has to learn from experience how to deal with many problems which arise in practice. This chapter is presented to point out some of these problems.

The eliciting of information from patients, which is the basis of psychological assessment, requires considerable skill and knowledge. Where testers are unskilled the consequences for both clinical and research work may be serious. In clinical work concerned with diagnosis, selection for treatment, follow-up of treatment results and vocational guidance the patient may not be helped in the best and most appropriate way. In research, the data may be incomplete if the required information cannot be elicited. Statistical compromises then have to be made to handle the incomplete data.

Experimental groups may be biased if some patients selected for study do not cooperate. This could happen, for instance, where psychopaths are being asked to undergo the Sedation Threshold Test (Claridge and Herrington, 1963) which involves venipuncture. Often these patients see no personal advantage in cooperating and refuse to do the test. Such drop-outs may seriously bias the group results. It is likely that the notorious difficulty in repeating experimental results in psychology and psychiatry is often due to differences in the level of skill of different researchers, some losing patients and data which more skilled experimenters can retain.

When the variables that may affect performance in psychological tests are recognized then some attempt can be made to control them. More valid data will be obtained if due regard is paid to the complexities of behaviour instead of oversimplifying test situations.

A person only becomes a skilled tester with practice. Experience is necessary before he can decide what is the best way to handle any particular situation. Sometimes one approach is appropriate whereas at other times, given what appears in many respects to be the same situation, the exact opposite approach is called for.

This view does not imply the creation of a mystique around testing nor that psychologists should be the only ones allowed to use tests. Demarcation disputes are alien to the spirit of the behavioural sciences, but all who use tests must be aware of their responsibility to acquaint themselves with the problems which testing involves.

Preliminaries and Introduction of the Patient to the Test Situation

Patients' attitudes to testing

When a patient comes to be interviewed by a psychologist he brings with him a set of attitudes and expectations derived from a variety of sources: from newspapers and T.V., from other patients, from nursing and medical staff and from his own previous experience.

One influence on patients' attitudes which is difficult to assess is the contact with other patients on the ward who have seen the psychologist. They undoubtedly discuss the procedures they undergo but ignorance of the effect of this contact tends to be rationalized by assuming it to be negligible.

The way in which he is told he is to see the psychologist will also influence a patient's attitudes. He may be informed by the psychiatrist, by a nurse or by the psychologist himself but whoever makes this communication can affect the patient's degree of cooperation and ego-involvement in testing. If he has not been prepared at all he may wonder who the psychologist is and even whether he is a member of the hospital staff. On the other hand, if he sees the psychologist about the ward each day and knows that other patients visit him, he may also expect to visit him and even feel left out of things if he does not.

Usually the doctor dealing with the patient will tell him that he will arrange for him to be seen by the psychologist. In doing this a doctor may convey the impression that the psychologist has magical powers and invest him with an aura of omnipotence. Generally though the doctor will make the referral in a matter of fact way as though the patient were to go and see the optician or have an x-ray.

When a patient comes to hospital he may expect to see a number of people who will question him. By the time he has seen the doctor at the out-patient clinic, the admitting doctor, a registrar who will handle his case, a consultant and a social worker, seeing the psychologist is just another in a long series of interviews. He may by this time be getting resistant to further interviewers or he may be warming to his role as a patient.

Preliminaries to testing

Before seeing a patient a psychologist will generally seek some information about him and his present condition from the doctor and nursing staff. They may be discouraging, for instance, in suggesting that 'you won't get anything out of him.' Indeed a doctor may misguidedly refrain from referring a patient because he is regarded as being in an untestable state. Usually one should not be reluctant to test a patient before actually making the attempt. It is often surprising how much cooperation can be achieved with attention to technique and some persistence.

The next stage in the procedure is to make sure that the patient fulfils the conditions required by the test situation, for example that he has glasses when he is required to read, that he is not stone deaf or that he has his hearing aid tuned when he is required to listen to a tone used to elicit a G.S.R., or that he is not colour blind if this is relevant.

Before testing it is necessary to establish a relationship which is conducive to the obtaining of the information required. A brief chat on the ward before the patient comes to the office may help. When he comes for interview a short conversation acclimatizes him and prepares his mental set for testing. Good rapport is not essential, although usually helpful. As long as he is ready to do what is asked there is no need to spend excessive time in developing a relationship.

In this conversation it is often worthwhile rapidly going over the patient's present condition and recent history so that the testing can be related to these. For instance, he might be asked how he is feeling, what does he feel is wrong with him, in what way does he feel he has changed, why has he come into hospital?

Introduction to testing

Having established contact in this way the testing can then be introduced. Sometimes there is no need to give any reason for this. One may merely say, 'I have some questions I'd like to ask you' (or 'tests I'd like you to do'). If this is said in a positive manner, assuming an attitude indicating that he is expected to cooperate then the patient is usually carried along without thought of objecting. At other times he may be prepared to admit, for instance, that his concentration is poor. Most psychiatric patients will confess that they are not able to concentrate as well as they used to. It is then but a short step to introduce the tests in order to see how he is concentrating. Depending on the tests to be given, one may get the patient to admit that his emotions sometimes get the better of him or his memory is poor or he finds it difficult to relax. The tests are then introduced as 'helping us to understand these difficulties better'. Even if he does not admit to such problems, or perhaps feels he has recovered from them, one can often carry on regardless by saying 'well we want to see how your memory [etc.] is now.'

In giving reasons to the patient for the necessity of testing one should be as truthful as possible, having regard to the patient's level of understanding. However, complete truthfulness may cause unnecessary suffering to the patient or make him reluctant to do the test. There could be little justification for saying to an intelligent professional worker, 'You are being tested so that we can assess the degree and quality of your intellectual deterioration in order to distinguish between the differential diagnoses of organic brain disease and schizophrenia.' One has to consider seriously, however, how far a patient should be misled concerning the reasons for testing. For example, in using projective techniques to assess emotional adjustment, to elicit conflicts and interpersonal attitudes it is questionable whether a patient should be led to believe his imaginative ability is being tested.

If the assessment is part of a research project there is the problem of asking his permission to carry out the procedures. Often there seems little point in doing this with psychological tests which are used in clinical work anyway. It is then probably best for the patient to regard the testing as part of the day-to-day clinical work. This is particularly true where it is eventually intended to use tests for clinical work. If they are developed under conditions different from those in which they are to be used in practice there may be important differences in results. For instance, in using G.S.R. habituation to assess anxiety (Lader and Wing, 1964) a patient may give differing results according to whether he is persuaded that he is doing the test to help the psychologist in his research or whether he feels his own future might depend on the outcome of the test. Where medical procedures are being carried out, as in the Sedation Threshold Test, permission has to be obtained from the patient, or if he is psychotic from a relative.

In this brief introductory period the psychologist has to gauge the patient's attitudes towards the interview and towards the psychologist. He may have to try to alter those attitudes so as to achieve a sufficient degree of cooperation from the patient in order to get the information he is seeking from the interview.

Persuading Patients to Cooperate

Most patients will cooperate in doing the various tasks that psychologists ask them to do. This follows partly from their curiosity and partly from the tradition of respect for authority figures. By his skill the psychologist attempts to develop a relationship and provide a situation in which the patient feels sufficiently at ease and secure to produce the information required. But some patients will object and the psychologist frequently has to deal with suspicion, fear and hostility before he can get such patients to cooperate.

There are various reasons why a patient may be unwilling to be tested. He may consider his condition to be physical and not mental. Therefore he cannot see the relevance of psychological tests. He may feel perfectly well and object to testing because this implies he is ill. He may think that the test results will be used to show that he is mad so that he can be locked up in hospital for the rest of his life.

A patient may refuse to see the psychologist when asked to do so by a nurse. The psychologist would then have to see the patient personally in the ward to tell him he would like to have a chat. Often this is sufficient and the patient comes willingly. After arriving at the interview and having been persuaded to talk about himself the tests can be introduced as a matter of course.

If he does not come he can be asked, 'What is the difficulty – what is worrying you – is there something on your mind?' He may mention fears about the interview and it may be fairly easy to relieve him of these. Suspicion and hostility can be allayed sometimes by reassurance and explanation. He can be told that the interview is for his own good and that it is necessary if he is to be helped. It may help to indicate that other patients also complete these tests. There are times when it is appropriate to ignore objections, gently taking the patient's arm and guiding him to the interviewing room insisting cheerfully that he really wants to come. Under such bustling encouragement many patients will not put up further opposition.

On the whole physical force should be avoided but there are occasions when it should be considered. For example, a restless hebephrenic girl who did a jig during testing rushed out of the room impulsively at one stage. When marched back again she was not in the least perturbed. There are occasions, however, when such action would be sheer folly. Everything depends on the mental state of the patient. It is probably advisable to leave these situations to the nursing staff who will often know best how to deal with a particular patient. Sometimes such a spontaneous interruption of the testing session is the means whereby the patient discharges his restlessness. It is then better to allow the patient who may have walked out of the room to pace up and down for a while before inviting him to start again. With chronic schizophrenic patients this is often all that is required.

Enthusiasm for getting the patient to cooperate carries the danger of illegitimate and unethical procedures, such as threats of withdrawal of privileges. It would be wrong to imply that discharge from hospital is dependent on cooperation. It is better to admit defeat rather than resort to such procedures. At such a time the nursing staff, the doctor or the patient's relatives may have suggestions for persuading the patient to cooperate. Alternatively a later opportunity may be taken to try again.

Testing Technique

In presenting test items a tester should take a positive attitude. This makes it difficult for the patient to refuse. It is better to say 'I want you to try this one now', rather than 'Would you mind doing this test?'

Checking the patient's understanding of instructions

Having secured his attention and cooperation the patient is given the test instructions so that he knows what he has to do. It is essential that his understanding of the instructions be checked. For example, in administering Piaget's tests of conservation (Inhelder and Piaget, 1958) to schizophrenic patients some patients will answer without really understanding what is asked. They may answer 'Yes' to every question. Thus, if shown a ball and a sausage, each made of the same amount of plasticine, and asked which has the more plasticine, the patient's answers can be checked by pointing to the ball and saying 'Is there more in this one than in this one?' To check again the sausage is indicated first and the question repeated. If he answers 'yes' to both questions he has evidently not really understood what is required of him.

In a distraction test the procedure is to present the task without distraction and then with distraction (Chapman and McGhie, 1962). It is not sufficient to tell him what he has to do in the condition of no distraction, however simple the task. One has to check that he really can do it and that he knows what is required. This will necessitate a preliminary trial run through and possibly repetition of the instructions or even a demonstration.

The use of gestures and miming is often useful in order to show a patient what is required. For example, in getting a senile patient with dementia to copy the Bender Gestalt Figures (Bender, 1938) one often has to place the pencil in his hand and guide it to the paper. When he wonders what he has to draw, the card in front of him is indicated. If he starts to draw on the design his hand has to be moved to the paper. By pointing to the card and the paper he is eventually brought to understand that he has to copy the design on the card on to the paper.

Procedure

In administering any test, of course, one should start out with the intention of adhering to the standard instructions given in the manual. However, there are frequent occasions when the standard procedure must be modified. For instance, a depressed patient who is retarded may take a long time over the Block Design Subtest of the W.A.I.S. (Wechsler, 1955). If he goes over the time limit it often pays to let him continue in order to see how much he can achieve when given unlimited time. This may be important where there is a question of organic involvement. The organic patient may also be retarded but may be expected to make errors however much time he is allowed.

A procedural modification which should only be used as a last resort is for the examiner to do a bit of a test which the patient will not begin. A patient may even say to the examiner 'I'd like to see you do it' when shown the T.A.T. cards. The only way to get such a patient to do the test may be for the examiner to make a start. This can happen with aphasic patients who are asked to say the alphabet. They may not be capable of starting until the first few letters have been said for them.

Sometimes it is not very clear whether a patient has finished doing a test such as the W.A.I.S. Block Design. It is then necessary to ask, 'Have you finished that one?' to find out whether he is still considering it or whether he is waiting for the next one. If he is not too sure about his production, this question may be taken by him as confirmation of his suspicion that he has not quite got it right. This apparently neutral question might, therefore, be taken by the patient as prompting. The tester must consider whatever he says and does very carefully to determine whether it constitutes prompting.

Prompting, reassurance and encouragement

Illegitimate hints may be given quite unwittingly by the tester as, for instance, in repeating digits. If the tester winces when he hears a wrong digit, the patient may be able to correct himself. Even nodding or saying 'Aha' to acknowledge the answers to Progressive Matrices (Raven, 1960) problems may be taken by a patient as indicating that his solution was correct. If his solution was based on a position response he may then repeat that position response for all the other problems, probably receiving confirmation of the correctness of the choice by a nod of the examiner's head.

When a patient asks for reassurance the tester should answer as non-committally as possible, but in such a way as to retain his interest and cooperation. If the patient asks, 'Should I do it this way?' the tester might repeat part of the instructions. In the Bender Gestalt, for example, he would be told to copy what he sees on the card. Or he might be told that he is to do it in his own way or that he is to decide himself.

Frequently patients ask 'Is that right?' In answering this question the examiner must refuse to inform the patient of the accuracy of his responses, but in such a way that his effort and interest will be maintained. When the response is made, however, and recorded, one may indicate whether it was correct or not. This can often be avoided by saying he is doing well. There is little point in saying he got an item right which, in fact, was wrong. If he realizes it was wrong he may not appreciate your attempts to jolly him along misleadingly. A better solution is to say, 'Well no, you didn't get that one. Have a go at the next one.' Honesty in this situation is often more likely to increase trust and cooperation than trying to cover up the patient's failures. If he realizes he is failing and appears to be discouraged he may well be told, 'Well they are meant to be difficult you know. You are not expected to get them all right.' Sometimes patients ask what is the correct answer. For example, many patients are curious to know what the objects in the W.A.I.S. Object Assembly are. Sometimes it does no harm to tell them, but if a patient is to be retested later the question should be parried because the knowledge may help him on retesting.

Patients often say 'I can't do it.' One answer to this is 'Do what you can, try it anyway.' Another way of answering is to be more forceful by insisting, 'Yes, you can do it, now tell me some more.' This would be the approach in the Object Classification Test (Payne, 1962) in which the patient has to be encouraged to continue sorting the first time he gives up. In testing for perseveration one test requires the patient to name all the animals he can think of (Freeman and Gathercole, 1966). When he gives up he should be vigorously encouraged to continue. Sometimes an item cannot be repeated, as in memory tests such as digit span, and if the patient says he cannot do it he can be told, 'never mind, try the next one.' In some tests he can be told to have a guess if he cannot do an item, as in the Progressive Matrices. Naturally it does not pay off to insist on the patient guessing if he feels strongly that he can only give answers that he is sure of. Many patients are affronted when asked to guess and if pressed too hard may refuse to finish the test.

Test order

When several tests are to be administered to a patient, the order in which they are given should be carefully considered because responses to later tests may be affected by earlier ones. This has been shown when several questionnaires have been given to patients (Furneaux and Lindahl, 1963). It is obviously necessary to give a test of G.S.R. habituation before a distraction test when auditory stimuli are used for both the G.S.R. and the distraction test. In correlating performance on the Epstein Test (1953) of 'overinclusive' thinking with a measure of intelligence such as in the Shipley Hartford or the Synonyms of the

Mill Hill Vocabulary (Raven, 1958b) there may be different results according to whether the Epstein is given before or after the intelligence tests. The reason is that in all three tests the subject has to select words which have to be underlined. The set for the Epstein Test acquired during the Mill Hill Vocabulary might be quite different from the set adopted without such prior experience. Another situation where order of administration is important is in testing for perseveration. If writing tests are given before drawing tests, the patient may continue to write when asked to draw.

The Influence of Patient Variables

In using psychological tests for assessing mentally disordered patients an examiner must be fully acquainted with psychiatric symptomatology in order to deal with the special problems of testing which they pose.

Anxiety

One of the main sources of interference with testing is anxiety. Sometimes a patient will tell you what is making him anxious. Sometimes he has to be asked. Occasionally one may be able to bring a worry to the surface for ventilation by verbalizing it for the patient in the same way as the psychoanalyst makes interpretations in psychotherapy. Such interpretations are sometimes right and appropriate and they result in considerable relief for the patient. Sometimes this does not happen and one has to be careful in presenting them so that the patient is not further alienated should the interpretation be inappropriate.

A common fear is of authority figures. Another is of ridicule and loss of approval if he fails to perform as he thinks he should. This fear may be increased when a stop watch is waved in front of him and he gets flustered as a result. Denial of loss of abilities sometimes occurs in brain damaged professional people and then the patient may not wish to be faced with his impairment. There may be sexual fears, as in a woman who became very tense when sitting at a table to do distraction tests whilst the tester stood behind her giving her the instructions. She became so distressed that she could not do the test and could not speak to explain her distress. Later it was found that she had been very upset in a crowded shop when a man had pressed up against her from behind and in the crush she had not been able to escape.

Some patients are afraid of being tested from their fear of going insane. The fact that they are being given tests which they think are used for mad people indicates to them that they are being considered mad.

There are innumerable delusional fears which can never be guessed at. For example, one woman was unable to press the key in a reaction time experiment because she thought she might thereby be torturing and killing her husband. In schizophrenic patients the disorder of identity and ego boundaries may be a source of distress. Many schizophrenic patients were reluctant to put on headphones in a room where a tape recorder was plain to see because they feared that their thoughts could be recorded on the tape. Some schizophrenic patients suffer from acute anxiety which may not have any special content. During this phase they may not be able to say what they are afraid of.

Patients are often afraid of what is going to happen to them. They may expect to die or be tortured or suffer in some way. This fear was raised by many patients coming to the psychologist's room which was in a unit of the hospital known as 'The Treatment Centre', where E.C.T. was given. A patient who had heard of E.C.T. and knew it was given in the Treatment Centre would understandably be worried if asked to place his hand on a metal plate with impressive wire connexions and electrical gadgetry attached. Then to place on his head headphones, which are very similar to the E.C.T. electrodes, is too much. This situation takes a lot of explaining to reassure fearful patients that all will be well. Some patients, usually those who are not afraid, are most grateful after this G.S.R. test, saying how much better they feel after the treatment!

Ideally one should use two rooms for such assessment procedures; the one for the patient being a pleasant, carpeted, soothing room, and the other next door housing all the equipment which can provoke such distress.

To relieve anxieties about what will happen the examiner should give a straightforward account of what will be done and what will happen to the patient. Even then there will be patients who are sceptical. After all, who can expect a psychologist to be honest and aboveboard when one has heard stories of experimental subjects who are tipped out of chairs without warning, who endure unexpected loud noises (Landis and Hunt, 1939) and who are led to think the psychologist is interested in one aspect of their behaviour when really he is interested in another (Bramel, 1962). It sometimes takes a lot of talking to convince patients you are a 'straight' psychologist.

There is a danger in over-explaining whilst attempting to reassure patients, that anxieties may be evoked which were not previously present. For instance, if whilst attaching G.S.R. electrodes a patient is assured confidently that he will not get a shock from this, the tester may be planting an idea that had not occurred to him. He is then more worried after reassurance than before.

Depression and elation

Apart from anxiety there are other disturbances of affect which have their influence on testing. Weeping and self-criticism frequently occur and usually call for sympathetic support. The euphoric patient may feel that there is no need for him to try and make an effort. The grandiose patient may regard the interview as an insult, particularly if he has a condition such as affect-laden paraphrenia. He may be humoured if the examiner pays deference to his delusional position.

Aggression

In general, verbal aggression can be ignored and testing can be continued without getting involved in an argument with the patient.

The patient who criticizes may be deliberately trying to provoke the examiner to anger. On the other hand, the schizophrenic who threatens to write to the authorities about the examiner's behaviour may merely be discharging emotion.

One way in which some patients may be made angry is by giving them tests which are ridiculously easy. In a study of the impairment of switching type of perseveration (Freeman and Gathercole, 1966), senile patients were asked first of all to count from one to ten and then to say the alphabet. It was not surprising that a number of them objected to being treated like children. However, when these two tasks are given after a number of more challenging ones in which patients have been engrossed, then very few will object to the easy ones, merely because they have become set for doing tasks.

With a patient who objects to the childish nature of the tests it may be worthwhile presenting a task he finds difficult. For instance, a patient who was given several psychomotor tasks objected to the peg-board as being infantile and it was only when he was given a complicated pencil maze to do with his non-dominant hand whilst viewing through a mirror, that he felt he was being treated like an adult.

When a patient objects to tests which he feels are more appropriate for babies, this may be a defensive manoeuvre to avoid failure. The examiner should be careful about showing the patient that he can fail in order to show him that they are not so childish as they seem. Such a patient will probably be afraid of the examiner's ridicule and the threat to his self-esteem if he should fail.

If a brain-damaged patient feels he is failing or that a problem is too difficult for him a catastrophic reaction may be precipitated (Goldstein, 1959). Apart from irritability, anger and refusal to cooperate further there may be a reduction in the level of consciousness.

Attention

In addition to the basic necessity of getting the patient to cooperate and do the tests, another basic requirement is that his attention should be retained in order to complete them. Since defects of attention are very common in psychiatric patients obviously the examiner has to give special consideration to these in order to determine in what way they are affecting test performance.

Confusion and perplexity are often features of psychosis, whether functional, organic or even drug-induced as by L.S.D. It is possible to increase this confusion if there are too many people questioning the patient at any one time. This can happen when a patient is being seen in bed and the nurse in attendance is trying to be helpful. She may repeat the question asked of the patient in an effort to get him to understand. Frequently this merely serves to confuse him all the more because he has to keep redirecting his attention to different sources of information. In such a case it is better for one person to do the questioning and encouraging.

Often psychotic patients are preoccupied or withdrawn and have to be alerted to external communications. One schizophrenic girl being given the W.A.I.S. vocabulary subtest answered with words concerning her delusions such as God, spirit, sky, etc. The whole list of words was presented a second time in a loud and staccato manner with a bang on the table to emphasis each question. This procedure enabled her to give quite rational and appropriate replies.

Distractibility and perseveration are symptoms of disordered attention which may interfere with testing and require special precautions for minimizing the interference. Distractions such as sounds outside the room and visual heterogeneity inside the room can be reduced. Then if the examiner is not restless and manages to keep still this helps the patient to attend to what he is saying. By giving instructions or asking questions slowly so that the patient has time to take in what is being said his chances of understanding are improved (Chapman and McGhie, 1963).

In dealing with perseveration of the impairment of switching variety the examiner should allow time between tests or questions and consider the order of presentation since some responses are more likely to perseverate than others.

In both perseveration and distractibility it is necessary to keep stimulus material out of sight. For example, in visual memory tests, the stimulus designs may have to be completely removed from sight and his previous productions also. Even other shapes and drawings which might be on the examiner's desk at the time should be carefully covered.

The first signs of fatigue in testing are usually irritability and disorders of attention. The examiner constantly has to be alert for these signs so that the patient may be given a rest or a change of task. Children and patients with organic cerebral disorder are probably most prone to fatigue but very often these patients are able to carry on for up to two hours without fatigue. Assessment should not therefore be ended prematurely, particularly where limited time is available as in out-patient consultations.

The Influence of Tester Variables

There are many aspects of the tester which will influence the patient's willingness to be tested and his attitude to the tests.

Appearance, sex, age, colour, accent, nationality are probably in themselves of lesser importance. Of greater importance are his adaptability, skill, tact, ability to empathize and his human interest in the patients he examines. As in any aspect of work with psychiatric patients he needs to have an understanding of his own personality functioning. This is necessary if his assessment of what happens in the relationship with the patient is to be as free from bias and distortion as possible.

He should be constantly on the look out for the effects of his actions and words on the patient being interviewed. Signs of impatience and boredom, such as yawning or looking at one's watch, are often taken by patients more seriously than we think. For example, one patient said when the examiner yawned, 'You must be tired – tired of this work, tired of patients, tired of me.'

Patients will often bring to the testing situation attitudes carried over from other situations in their lives such as attitudes towards members of the family or to other members of the hospital staff. The psychologist therefore may not be treated as a person in his own right but as a representative of a class of people for that patient. This was apparent in the case of the patient with the sexual fears described above.

Although the psychologist cannot always know the origins of such attitudes within the test situation he should be aware that they occur and be prepared to handle them when he recognizes them.

The biasing influence of testers on results of experiments has been examined recently rather more than in the past. The various biases that can occur have been called 'the experimenter effect'. One aspect of this effect is the personality of the tester. Another is his attitude or orientation to the results of the experiment or test.

Outcome orientation

A tester can bias a patient's behaviour and responses to tests as a consequence of his expectations of the way the patient should behave. 'One prophesies an event and the expectation of the event then changes the prophet's behaviour in such a way as to make the predicted event more likely' (Rosenthal, 1964, p. 405). This is the self-fulfilling prophecy. The behaviour of the tester in some way alters the subject's responses to fit in with the behaviour expected. Testers are often not as impartial as they would like to think of themselves. They often have a personal orientation to the outcome of their tests and the results may give rise to varying emotions depending on whether the tester's expectations and wishes are fulfilled or not.

The explanation of the effect of the tester's outcome orientation probably lies in the unintentional transmission of sensory cues from tester to patient. There was a remarkable horse called Clever Hans, which could count by tapping his hoof on the ground and perform various other remarkable tasks. It was found that he was using visual cues transmitted by his questioners. A person asking Hans to count would incline his head forwards when he expected him to begin counting. When he expected Hans to stop, a questioner would slightly raise his head or eyebrows.

In studies where subjects are handled by their experimenters tactile and kinaesthetic cues are probably important. Rats which were actually selected randomly but said to be rather good at running mazes (maze bright) were found by their experimenters to be maze bright. Rats said to be maze dull (but also in fact randomly selected) were found to be maze dull. The performance of these rats in the maze was in some way affected by their experimenters' expectations.

In clinical testing patients receive abundant cues both visual and auditory from which to judge their tester's expectations. The tone of voice, inflexion and the whole manner of speaking of the tester conveys a lot about his attitude to the patient's responses. Apart from consciously directing the patient's responses he may verbally condition them unintentionally. It has been shown that it is possible to increase the subject's use of plural nouns by saying 'mm-hmm' whenever such words are used by the subject. This effect can even occur in a multiple choice test such as the Progressive Matrices. A patient quickly developed a position effect, always selecting the item in the bottom left-hand corner, as a result of the author's saying 'mm-hmm' when the patient gave each answer. Obviously such conditioning could affect projective test responses as well. Choice of T.A.T. content might well be affected by the reinforcement provided by the tester.

Visual cues are provided by a tester's facial expression, his gestures and general bodily disposition.

There must be individual differences in the extent to which different people are able to pick up auditory and visual cues. It may be that some subjects are so insensitive that whatever the tester does to indicate his orientation to the outcome of the tests has no effect on their results. Other subjects may be hypersensitive so that they respond to too many cues and misinterpret the tester's expectations. These aspects of the experimenter-subject relationship have barely been touched in research yet.

One of the points in testing at which the tester has to be most wary of providing cues for patients is when elucidation of a response is required. In the W.A.I.S. Comprehension subtest the tester may have to ask the patient to explain his initial response to a question. In doing this he may unwittingly vary his questioning in such a way as to elicit behaviour which will confirm his expectations about the patient.

In open-ended tests, where judgements have to be made in scoring, a tester may score harshly in one case but leniently in another. This may be done quite unwittingly but with the effect of confirming expectations.

A tester might derive his expectations as to what the patient will do, in different ways. He may make up his mind about the patient after reading the case history or hearing the psychiatrist's opinion. If he has no definite expectations before he starts then he may be influenced by the early responses in the test session. These are known as early data returns. The tester may perceive some apparent trends in the early responses which lead him to influence the ensuing responses so as to confirm these trends. Rorschach users are even advised to start assessing responses from the beginning of the administration of the cards. It is suggested that this procedure heightens the alertness of the tester in what otherwise might prove a rather tedious task. Obviously the early data returns may well provide expectations as to what kind of responses the patient will make later.

Although the effects which testers can have on results have long been recognized the interest shown in this problem has been negligible. Few psychologists have recognized it and those who have have mainly discounted it. Recently there has been an upsurge of interest and a number of reviews have appeared (Kintz *et al.*, 1965; Rosenthal, 1964). Undoubtedly this neglected variable, the tester, will be much investigated in future. With more knowledge it will be possible to measure the extent of the influence of the tester on the results he obtains and some degree of control over this influence may eventually be gained.

In the meantime what steps can be taken to minimize the effect of the tester's outcome orientation? One manoeuvre is to have tests administered by testers who know nothing of the patients they test, and who are not told the reasons why they are giving the tests. The patient is seen by one psychologist who decides what tests will be used. He is then handed on to another psychologist or a testing assistant who administers the tests. Testing is thus 'blind' in the same sense that drugs are administered 'blind' in drug trials. Unfortunately there is evidence that even this procedure does not eliminate the effect of the tester. In one investigation it was found that the experimenters' expectations can be subtly communicated to the testing assistants. The only trouble with experiments on the experimenter effect of course is that they themselves are open to bias from the experimenter effect!

Another manoeuvre is for the tester to present the test record for analysis to an interpreter who knows nothing more than perhaps the patient's age and sex. In this case the interpretation is done blind, since he does not know the reason why the test is given. If both of these manoeuvres are adopted, testing blind and interpreting blind, then a team of three would be required for the assessment of each patient: the person who decides what tests shall be given, the tester, and the test interpreter.

The influence of the tester on test results is minimized in tests that are self-administered. Where the patient can read instructions and do a test by himself, interaction between patient and tester is virtually eliminated. It may even be possible to devise machines, programmed to instruct the patient what to do. This in fact has been done on a limited scale already. The patient's understanding would be tested on examples provided in the programme and he would only be introduced by the machine to the test proper when he had grasped what he had to do. This may work where the subject's motivation and understanding of the tasks set him are adequate. In psychiatric work however this is frequently not the case. Patients often have to be persuaded to do tests and then encouraged to continue to the end. If the patient has anxieties about the tests these can be discussed with a tester but not with a machine. The reverse is also true. Many patients are much happier doing tests alone than with a tester hovering over them.

One way of eliminating the transmission of visual cues from tester to patient in the test session would be to place a screen between them. Space could be left between the table top and the bottom of the screen to allow the tester to see what the patient is doing and to present test materials to him. This arrangement would naturally eliminate the transmission of visual cues from the patient to the tester so that the patient's grimaces, boredom, excitement, fear and so forth would be lost to the tester.

Perhaps the most significant implication of this discussion for testing is that testers should be alerted to the experimenter effect so that even though they may not be able to measure its effect nor control it, at least they may allow for it in drawing conclusions from test data. Undoubtedly the first step on the way to controlling this effect is to recognize that it exists. It cannot be long before this recognition becomes universal among psychologists. Only when it is recognized can further steps be taken towards controlling it.

How important in clinical psychology is the experimenter effect? Certainly it has been demonstrated in a number of studies to have a significant influence in a variety of experimental and test situations. But does it affect the actual conclusions drawn from tests in individual cases? It may be that the effect is slight and only achieves significance when large numbers of subjects are used. If this were the case it could be safely ignored in considering the results of the individual patient. Probably it is significant in some cases but not in others. At the present state of our knowledge it is safest to accept that it can have appreciable effects and to regard it as potentially biasing in every case.

The effect can enter into assessment whichever method is used. When statistical methods of inference are used it is less likely to obtrude at the level of inference. At the level of testing, the effect is as likely to occur here as when other methods are used. When clinical inference is used it could create havoc if it were not recognized. The psychologist who is aware of the effect may be able to make some allowance for it in drawing his conclusions. Oddly enough if it goes unrecognized, it is as likely to bias findings when experimental investigation of patients is undertaken as when either of the other two methods are adopted. If it *is* recognized then the experiments can be designed to take account of the experimenter effect.

To summarize: It is known that the clinical psychologist can influence the sort of test results he gets from patients he tests. Often the extent of this influence is not known. However, there are various steps which can be taken to reduce this influence. Although there may be objections to these attempts to eliminate the effect they should nevertheless be taken to their logical conclusion in order to discover what their limitations are in practice. It may turn out ultimately that it will be possible to introduce the patient to a machine which will administer tests, score them, analyse the results and print out its conclusion! (See chapter 8.) Gedye (1966) has already demonstrated that this can be done.

8 Research in Assessment

This book has been mainly orientated toward the practical contribution which clinical psychology has to make to psychiatry. Consequently in considering research in assessment the aim here is to draw attention to research needs at the applied level. Although research having less immediate relevance may ultimately prove to have a great influence on day-to-day practice, there is so much bread-and-butter research that needs to be done that the former will not be discussed further here.

The Need for Research

In the first chapter the need for psychological tests was demonstrated and their advantages were outlined. In subsequent chapters attention has been repeatedly drawn to their shortcomings and limitations. The shortcomings listed below are brought together to show how vulnerable to criticism are the tools of the clinical psychologist. If these criticisms are to be overcome, extensive research is urgently required.

(1) *Subjectivity* Despite the inexorable movement towards objectivity many tests are still used subjectively. The extensive use of the clinical method of evaluating tests is not matched by the extensive validation of clinicians' judgements.

(2) *Triviality* The emphasis on objectivity has led to the development of many tests which are certainly objective but which produce such trivial findings as to be of no practical use.

(3) The *validity* of many tests leaves much to be desired. One reason for this is that psychiatric diagnosis has generally been the criterion for validation of tests. This is

not a satisfactory criterion because of the lack of agreement between psychiatrists in diagnosing patients.

(4) *Diagnostic implications* Very few tests have been validated against treatment-outcome and follow-up so that the practical implications of test results for prognosis and treatment are generally unknown.

(5) *Descriptive implications* Frequently it is not clear what psychological functions are being measured by a test.

(6) *Base rates* For many purposes it is essential to know the base rates if tests are to be used. Adequate data are in practice rarely available.

(7) *Overlap* Interpretation of scores of individual patients on tests developed with groups of subjects is often impossible because overlap of scores of groups is so great.

(8) *The experimenter effect* may influence results on any test. The extent of this influence is virtually unknown.

(9) *Standardization* Tests are frequently not standardized on the type of patients for whom they are intended in clinical use. Thus a test of thinking disorder may be shown to discriminate between patently thought-disordered schizophrenics and normals but it is the not-so-obvious case which is referred to the psychologist.

When listed in this way the cumulative effect of these criticisms is rather devastating. It should be stressed that *not all* tests are subject to all of the criticisms. One also takes comfort from the criticisms of psychiatric assessment made in the first chapter to justify the use of psychological tests. If the psychologist is not a good oracle at least he may well be better than alternative oracles. More important he is a self-correcting oracle and more alert to the need to carry out research to improve his efficiency.

The Need for Clinically Based Research

The great need for clinical psychologists to do research does not mean that they should retire from the clinical scene only to return when they have performed the necessary research. On the contrary, research in clinical psychology has to be pursued in a clinical context. Where it becomes divorced from clinical practice its clinical relevance diminishes. Researchers must inform themselves by clinical experience. Clinical psychology has been a child of whom people have had great expectations; only by the constant interaction of research and practice can it fulfil that promise.

In arguing this point one is a little embarrassed by the fact that much of the clinically based research in clinical psychology in the past has proved rather unfruitful. It is true to say that recent progress has been made by psychologists who have not been committed to clinical work, because they have been able to take an independent and detached look at the field. But it is painfully obvious when research is done which is not informed by clinical knowledge and insight. We are now in a position to go back to *clinical* research, enlightened by this visit to the laboratory.

It is the combination of clinical experience and research know-how that is powerful. Clinically based research may be practically orientated but frequently makes perfectly valid scientific contributions. It is in just this kind of area that distinctions between pure and applied science become meaningless. An example of the kind of research which is based on extensive clinical knowledge is the work of McGhie and his colleagues (McGhie and Chapman, 1961; Chapman and McGhie, 1962). After careful description of the subjective experiences of early schizophrenics they attempted to provide standard laboratory conditions to observe the disordered attention which they held to underlie the symptoms described by their patients. This laboratory simulation of the conditions which provoke symptoms and the attempt to measure abnormal behaviour in this way is a valuable step in the direction of achieving experimental control over symptoms. On the basis of their clinical and experimental studies they have suggested ways in which communication with schizophrenics may be improved. This involves limiting the input of information to a level at which the patient's disordered attention mechanisms can cope. This work, then, has therapeutic implications as well as advancing our scientific knowledge of schizophrenic behaviour.

Another way of formulating the problem of clinically based research is to say that one can find rigorous research (meaning good design, use of controls where necessary, use of objective techniques, sophisticated and appropriate statistical analysis) applied to significant problems as well as to trivial problems. One can also find sloppy research on both significant and trivial problems. There are arguments as to whether it is preferable to have sloppy research on significant problems or rigorous research on trivial ones. Ideally rigorous research on significant issues is to be sought but the achievement of this depends on the training of the investigator. Those who have invested time in getting training in research techniques often lack the knowledge to pick significant problems so they fall back on laboratory-generated problems which take them away from the practical problems of living which patients have to face. Such researchers consider they are being true to their scientific integrity in doing fundamental work. Those who have invested time in getting clinical experience and are more aware of significant issues tend to be lacking in the art of research. Although these two stereotypes exist there are enough clinical psychologists who manage to combine the two disciplines to demonstrate that good research into significant problems can be done.

Individual Cases

The work of the clinical psychologist provides him with the opportunity of studying individual patients in depth. The emphasis on group statistics and nomothetic research has led to a neglect of the scientific study of individuals. Yet there is a very respectable history, in psychology and medicine, of significant work resulting from single cases. Ebbinghaus' study of memory using himself as subject is a classic example.

There has been no-one whose influence on both psychology and psychiatry has been greater than Freud and Pavlov. Yet the works of both men were based on intensive study of single subjects: in Freud's case, patients, in Pavlov's case, dogs.

The detailed study in depth of a few patients is often more productive than superficial studies of large numbers of people. Idiographic and nomothetic approaches are of course complementary. It is a pity that the idiographic has been so much neglected in clinical psychology.

One way in which a commitment to the study of single cases benefits the clinical psychologist is that he is prepared to make the best research use of infrequent conditions. He becomes serendipitous. If something happens relating to the behaviour of a patient, which is unusual but which the psychologist recognizes as having important implications for either theory or practice, he should be prepared to investigate this to establish just how important these implications are. By acquiring techniques and methods for studying single cases the psychologist is improving his powers of serendipity.

In the study of single cases for research purposes the experimental method described in chapter 3 is the method *par excellence*. Research having implications beyond the individual patient has frequently developed from the detailed investigation of single patients initially carried out for purely clinical purposes.

An important research task having a clinical orientation is to find out how far tests, which have been developed in research with groups of patients, can be applied to individual patients.

In research a commonly used method of achieving positive results is to make up two groups from subjects located at the opposite extremes of some dimension and to compare them on a test. The sedation threshold distinguished satisfactorily between extremes of dysthymics and hysterics and a claim was made that it was useful in the clinical investigation of patients (Claridge and Herrington, 1963). When Moffat and Levine (1964) examined this claim by testing a whole range of neurotic and character-disordered patients including the middle range, they were unable to confirm it. To test the application of research findings in the clinical situation requires a research project in itself. The aim of such studies is to bridge the gap between research (in which the primary aim is not clinical) and the clinical application of those research findings. Even when such studies are negative they may expose variables which require further study before positive clinical relevance can be established. The practically orientated research then feeds back to the so-called pure research.

Basic Data Collection

If nomothetic tests are to be more useful in individual cases a great deal of basic information about those tests will have to be collected. This is a continuing need. Standardization for local use of tests standardized elsewhere is necessary. The use of the M.M.P.I. and W.A.I.S. in Britain has always been questionable without local norms. Local base rates will need to be sought for decisions as to the utility of tests. If the diagnostic implications of tests for treatment, prognosis, description and etiology are to be known, norms for each of these categories will be required. Test data from which probabilities attaching to scores of individual patients can be determined will be necessary.

There is a great need for more comparative data from normals. Little is known concerning the relative rates of decline in short-term memory for visual and auditory material in the normal process of ageing. Data on normals is essential for the assessment of pathological memory decline.

Automation in Assessment

The introduction of automation into clinical psychology can fairly be said to have hardly begun. But enough has been done to indicate that this is going to be one of the growth points for far-reaching future developments.

The same objections have been raised to automating assessment procedures as have been raised to automating teaching by means of machines. It is suggested that the essentially human aspect of psychological assessment is lost or that psychologists will become machine operators. The answer to these points is that individual testing by psychologists is a very time-consuming business and the supply of clinical psychologists will not meet the need for the foreseeable future. We will have to make use of whatever technical aids we can if an adequate psychological service is to be provided. If automation can take over some of the functions of the psychologist he is thereby released to concentrate on the tasks for which he is essential, one of which is to develop improved methods of assessment.

The objection to the dehumanizing of assessment is discussed in chapter 7 where it is shown that testers can bias the results which they get from tests and it is suggested that automation may be positively useful in overcoming such distortions.

Machine scoring

Many tests such as questionnaires require a patient to leave a record of his responses as he does the test. With the advent of mark-sensing devices, record sheets can be processed by a machine, which gives a print-out of the patient's test results. This can provide an enormous saving of the psychologist's time.

The record sheets have to be specially prepared for this purpose and generally the patient fills them in with a thick pencil by marking the paper with a line in the appropriate place.

Analysis of patient's responses

Computers have been programmed to take over the detailed statistical work which can be involved in analysing the test results of individual patients. P. Slater (1965) has prepared a program for the analysis of repertory grid data involving the calculation of many rank-correlation coefficients and a principal components analysis.

Where analysis calls for calculations of the kind recommended by Payne and Jones (1957), or reference to normative data concerning base rates or the application of objective rules such as Meehl and Dahlstrom (1960) describe, then these procedures can be built into computer programmes for routine use.

Beyond straight calculation is the detailed interpretation of test results. Finney (1966) amongst others, has described the programmed interpretation of questionnaire data. The method 'selects for each individual about a hundred statements from a repertory of about a thousand. It combines the separately chosen statements into a report reading much like a spontaneously composed one. It attempts to describe the conscious or unconscious motives and defences and the patterns of interpersonal relations typical of each person' (p. 81). All this is based on the patient's M.M.P.I. and C.P.I. protocols fed into the computer which out-puts a report as described above.

Administration of tests by machine

Many tests such as questionnaires and cognitive tests administered to groups of subjects are virtually self-administering. These are relatively easily adapted to machine administration.

Where testing requires a patient to respond to sequences of stimuli, as in the measurement of reaction time, the presentation of stimuli and the recording of responses can be automated to relieve the tester of the job of pressing keys, reading dials, noting scores and so on.

The new technology of teaching machines has been adapted by Gedye (1966) to the automation of procedures for testing learning and memory. But Gedye's configuration (now in commercial production) will undoubtedly have very general applications and recommends itself to the whole field of assessment in clinical psychology.

Multi-access, on-line, real-time computer facilities

The computer world is providing us with a whole new set of jargon to juggle with. 'Multi-access' means that there is more than one terminal to a computer. With very large computers two hundred people or more may be in simultaneous communication with one computer. Each user may have his own in-put and out-put device in his office. This of course poses problems of security for confidential records and various methods have been devised to ensure that unauthorized people cannot elicit information.

'On-line' means that a device is directly connected to a computer. 'Real-time' means that the computer can immediately deal with the in-put from various sources so that a user does not have to wait for one job to be completed before he can use it.

The advances in computer technology mean that a hospital's own computer can be linked with various departments, each of which will be able to use the computer as the need arises. Data can be fed in and out-put as required. Thus the pathology laboratory can feed in results of tests or even have measures actually being recorded on the patient fed directly to the computer. Nursing reports, doctors' reports, and social workers' reports can be in-put as soon as they are available. Similarly psychologists' test results and reports can be in-put as they become available. Indeed where psychological tests are being administered by teaching machine they also can be linked directly with the computer.

The form of the out-put will be determined by the needs of the questioner. A clinical psychiatrist may ask for a complete case history, a summary, or the results of a particular test. As Finney has shown, computers can be programmed to arrange their out-put in readable and sensible English. An administrator may want to find out how many patients of a particular kind the hospital is dealing with over a period. The psychologist may wish to find out the base rates for some condition.

One of the great advantages of computers to the hospital service is that they provide a means of keeping permanent and easily accessible records. Indeed information retrieval is so rapid that analysis of records is now possible which previously could not be contemplated.

A hospital computer therefore serves clinical, administrative and research needs.

Current applications and future developments

Each one of the automated methods described in this chapter has actually been put into operation. This is not a fantasy of what might happen in future but a statement of what has already been done.

It would be wrong to give the impression that these methods are widely available in this country at the present time. Nearly all regional hospital boards in Britain have computers but none provides multi-access facilities yet and there is virtually no computer time being provided for clinical applications. At the time of writing one or two hospitals in this country have small computers for clinical work but no hospital is providing computer time for the kind of clinical psychological and psychiatric applications described in the preceding sections.

Before automated methods are introduced on a wide scale considerable technological development will be required, but as they become more available so they provoke more research. Yet again clinical psychology is seen as both applied and pure science.

Role of Theory

There is an undoubted need for a self-conscious appraisal on the part of researchers in this field of the need for theory.

Tolor and Schulberg (1963) point out how so-called empirical research is often of little value. In evaluating research on the Bender–Gestalt Test they note the 'great preponderance of low level conceptualizations' and the need for more 'theoretical underpinning for the Bender–Gestalt Test which would clarify exactly what the test measures and how it does so' (p. 202).

There are two sides to this problem. On the one hand there is the spectacle of speculation running far ahead of research findings. Theorists are reluctant to wait until findings have been established, replicated, crossvalidated and consolidated. While on the other hand research is meaningless unless and until it is related to theory.

Exploratory research, in which the researcher says 'It would be interesting to see what would happen if we did this', tends to be regarded with some contempt. But this sort of approach is a necessary stage in the development of any science. However, too often as Bender (1963) points out exploratory research is undertaken when hypotheses are available for testing but are ignored.

Generally there should be little difficulty in relating research to theory considering the number of contending theorists in the field. Some theories are concerned only with specific behaviours, having little general application beyond that behaviour. Others are very general having implications for apparently quite unrelated areas of behaviour.

Some researchers attempt to synthesize different approaches and take what they can from wherever they can get it. Others prefer to follow a particular line with single-mindedness to see how far it will take them.

If the complexity of human behaviour is recognized then it is clear that no one approach will provide all the answers and that in the present state of our knowledge diversity of theoretical approaches is to be encouraged.

Appendix 1
Descriptions of Selected Tests

The following notes have been included for the benefit of readers who are not familiar with the tests mentioned in the text. Although many of the tests in this list are widely used many are not. Indeed many are outdated or for research only. The list should not be taken as a representative sample of tests used in clinical psychology. A fuller list of tests used by clinical psychologists is given in appendix 2.

For further description and evaluation of tests the reader is referred to the manuals on the tests, to Buros (1965) and to the journals.

Tests which are commercially available can mostly be obtained either from the National Foundation for Educational Research, The Mere, Upton, Slough, Buckinghamshire, England, or from the Psychological Corporation, 304 East 45th Street, New York 17, New York, U.S.A., both of which provide catalogues of tests.

A.H. 5 (Heim, 1956)

A test of high level intelligence in two parts: Part 1 requires verbal and numerical skills, Part 2 requires visual analytical skills. Sometimes used in university selection. British norms.

Archimedes Spiral (see reviews in Buros, 1965)

A spiral drawn on a card which when rotated gives an illusion of movement, either expanding or contracting. When a person has fixated this for a period and the card is stopped the person experiences an after-effect of apparent movement in the opposite direction. The length of time this after-effect takes to wear off varies and has been related to psychiatric variables such as brain disorder, hysteria-dysthymia and psychosis.

Bender's Visual-Motor Gestalt Test (Bender, 1938)

Nine designs copied one at a time by the patient. Impaired patients tend to make more errors. Originally protocols were assessed qualitatively for breakdown of the gestalts. Scoring systems are now available (e.g. Pascal and Suttell, 1951).

Bernreuter Personality Inventory (see reviews in Buros, 1965)

A pre-war questionnaire claimed to measure introversion, neurotic tendency, social dominance, confidence, self-sufficiency and sociability but discredited by factor analytic work and superseded by other questionnaires.

Body-Sway Test of Suggestibility (Eysenck, 1947)

The patient stands upright and listens to a record telling him he is falling forwards. The extent of body movement is measured.

California Psychological Inventory – C.P.I. (Gough, 1957)

Personality questionnaire with eighteen scales concerning social qualities such as tolerance and responsibility.

Cattell's Sixteen Personality Factor Questionnaire – 16 P.F. (Cattell and Eber, 1957)

Based on extensive factor analytic studies and providing scores for sixteen independent factors, as well as two second order factors – Anxiety and Extraversion, which are virtually identical with Eysenck's N and E scales (see

E.P.I.). Two parallel forms (A and B) and a shorter form (C) in simplified English. Each form takes about fifty minutes to complete.

Coloured Progressive Matrices (Raven, 1958a)

An adaptation of the Progressive Matrices for use with children, subnormals and old people.

Distraction Tests (Chapman and McGhie, 1962)

Devised in an attempt to measure distractibility. Patients are required to perform various tasks such as tracking the lateral direction of movement of a spot of light whilst being subjected to distracting irrelevant stimulation such as sounds presented over headphones. The errors in the performance of the task, with and without distraction are compared.

Epstein's Test of Overinclusive Thinking (Epstein, 1953)

A paper and pencil test of schizophrenic thought disorder in which five words are presented, alongside a word designating some concept such as 'house'. The patient has to underline words which are a necessary part of that concept, e.g. walls and roof are, but telephone is not. Overincluders tend to underline many more unnecessary things. Not adequately developed for clinical use.

Eysenck Personality Inventory – E.P.I. (Eysenck and Eysenck, 1964)

This revision of the Maudsley Personality Inventory (M.P.I.) provides scores for two scales – Neuroticism (N) and Extraversion (E), as well as a lie scale. Extensive British norms. Takes 5–10 minutes to complete.

Family Relations Indicator – F.R.I. (Howells and Lickorish, 1962)

Thirty-three cards with drawings of family situations. Used with children who are encouraged to make statements about the activities of the people shown, rather than make up stories. Used to investigate the child's relationships with other members of the family.

Galvanic Skin Response – G.S.R. (Venables and Martin, 1967)

Also called psychogalvanic response (P.G.R.), electrodermal response (E.D.R.). A physiological measure of palmar sweating which is one component of emotional arousal. Often measured together with other arousal components such as muscle tension, respiration, blood volume, pulse rate, pupil size, etc., by means of a polygraph. Popularly known as a lie detector.

Goldstein–Scheerer Tests of Abstract Ability (Goldstein and Scheerer, 1941)

A battery of five tests requiring visual and motor abilities: cube test, colour sorting, object sorting, form sorting, stick test. Used in assessing abstract–concrete attitudes in brain disordered patients. No quantification therefore no published standardization.

Level of Aspiration (Eysenck, 1952)

The patient performs some task, such as tracking a wavy line on a revolving drum. He makes an estimate of how well he will do next time and the task is repeated, after which he says how well he thinks he has done. This procedure is repeated several times. Scores are derived indicating (1) the average discrepancy between his actual score and his estimate of how well he would do; (2) the average discrepancy between his actual score on the present trial and his estimate for the next one.

Memory for Designs – M.F.D. (Graham and Kendall, 1960)

A series of fifteen designs presented briefly one at a time to the patient who has to draw each design on paper as soon as it is removed. A test of visual retention which takes about fifteen minutes. The quality of reproduction is often impaired in brain-disordered patients.

Mill Hill Vocabulary Test (Raven, 1958b)

Includes a multiple choice (synonyms) test and a definitions test of word knowledge. British norms. Takes about ten minutes per patient. Used in conjunction with Progressive Matrices.

Minnesota Multiphasic Personality Inventory – M.M.P.I. (Hathaway and McKinley, 1951)

Five hundred and fifty statements answered 'true', 'false' or 'cannot say'. Provides scores for nine clinical scales, but research has extended this (see Taylor's M.A.S. for example) to over two hundred scales of ego strength, hostility, etc. One of the most popular questionnaires.

Modified New Word Learning Test – M.N.W.L.T. (Walton and Black, 1959)

The meanings of the first ten words on a vocabulary test which are failed by a patient are taught him. After each presentation of the list of meanings the patient is tested to see how many he has learned. Word learning tends to be impaired in brain-disordered patients.

Nufferno Speed and Level Tests (Furneaux, 1960)

Problems of the letter series type. Level test allows unlimited time for patient to show how far he can get without a time limit. In the speed test his rate of work on problems he gets right is measured. British norms.

Objective-Analytic Personality Test Battery (Cattell *et al.* 1955)

A number of tests including reaction time, a finger maze and galvanic skin response. Some requiring paper and pencil, some apparatus. Time-consuming to administer. Used to measure eighteen traits such as 'general inhibition'.

Peg-Board (as in G.A.T.B. – see United States Employment Service, 1966)

A motor task in which the patient has to transfer pegs which fill the holes in one half of a board to the holes in the other half which are empty. The rate at which the pegs are transferred is a measure of his dexterity.

Piaget's Conservation Tests (Piaget, 1929)

Used with children to determine conceptual development. In one test two equal amounts of the same material, for example, plasticine, are made into different shapes – ball and sausage, for example – and up to a certain stage the child conceives of the amount as varying with the shape, even when he sees that none has been added or taken away from the original amounts.

Porteus Mazes (Porteus, 1965)

A series of mazes increasing in level of difficulty. Patient has to trace his way with a pencil through each maze which is printed on paper, without crossing lines, going down blind alleys, lifting his pencil, etc.

Two scores are derived: (1) test age which is regarded as a measure of foresight and planning, (2) an error score

(Q score) summarizing the number of times pencil lifted, etc., which has been related to temperament and delinquency.

Progressive Matrices (Raven, 1960)

A multiple choice test of perceptual reasoning which is relatively uninfluenced by educational and cultural background and therefore is a good indicator of native ability. Takes about forty to fifty minutes per patient. British norms. Widely used. Virtually self-administering. Used in conjunction with the Mill Hill Vocabulary Test.

Q Sort (Block, 1961)

Statements on cards are sorted by the patient according to some direction, for example, from 'most like me' to 'least like me' in a fixed number of categories (say nine) and the number of cards to be placed in each category is fixed to give a near normal distribution. A second sort for 'like I'd like to be' would provide data which could be correlated with the first sort. Some standard decks of cards are available. Often *ad hoc* cards are used.

Reaction-Time – R. T. (Woodworth and Schlosberg, 1954)

The patient holds a morse key down and is instructed to release it when he receives the signal which may be a light flashing on. The time taken to react to the signal is measured by an electronic timer.

Repertory Grid Technique (Bannister, 1965)

A method by which the constructs (attitudes, sentiments, concepts) of a patient can be studied. Many variations in content and scoring, but the basic method is to get the patient to apply a set of constructs (e.g. kind, honest) to a set of elements (e.g. people he knows, photographs of people). This is often done by ranking the elements. The interrelationships between the constructs are indicated by correlations.

Rorschach Ink Blots (Klopfer and Davidson, 1962)

Ten cards with ink blots shown one at a time to the patient who is asked to say what they remind him of. In structuring such unstructured material as ink blots people are assumed to 'project' their own personalities. This is therefore a projective test. Responses are classified and related to personality traits and diagnostic categories. Very popular with clinical psychologists in America, less so in Britain.

Sedation Threshold (Claridge and Herrington, 1963)

A measure of tolerance of a depressant drug. Sodium amytal is injected into the patient at a known rate whilst he doubles digits presented to him every two seconds. The point at which he begins to make errors is his sedation threshold and has been related to many psychiatric variables especially hysteria-dysthymia.

Semantic Differential – S.D. (Osgood *et al.*, 1957)

The patient rates various concepts on a number of rating scales, such as good–bad, strong–weak. Concepts selected depend on the assessment problem posed, but might include the patient's symptoms or socially desirable behaviour. Scores may be derived for the three major factors demonstrated by factor analysis – evaluation, potency and activity and the concepts plotted graphically in 'semantic' space.

Shapiro's Personal Questionnaire (Shapiro, 1961)

Statements elicited from a patient about his symptoms are scaled for severity. A paired comparison technique is then used to provide scores for symptom groups. When the questionnaire is administered over a period of time the relationships between the scores show how the symptoms are changing relative to each other. Used particularly in assessing the effects of treatments.

Shipley Hartford Scale (Shipley, 1946)

A brief screening test with two sections: (1) vocabulary, (2) reasoning. A discrepancy between scores on the two sections was intended as a measure of cognitive impairment. Now occasionally used as a quick test of intelligence.

Taylor's Manifest Anxiety Scale – M.A.S. (Taylor, 1953)

Fifty-item questionnaire. Items taken from M.M.P.I. Highly correlated with E.P.I. N scale and 16 P.F. Anxiety factor.

Thematic Apperception Test – T.A.T. (Murray, 1943)

A series of cards showing scenes, around which the patient is asked to make up stories. Each picture allows of a range of different interpretations, and it is assumed that the story selected by any particular person is determined by his personality. In the analysis of the stories attention is paid to the themes, needs and pressures expressed.

Wechsler Adult Intelligence Scale – W.A.I.S. (Wechsler, 1955)

Widely used by clinical psychologists. Requiring individual administration. Takes about 1¼ hours to test one patient. Well standardized in America, but no British norms available – eleven subtests, covering a wide range of abilities, are grouped into two scales: (1) verbal, including vocabulary, comprehension, (2) performance, including manipulative and perceptual tasks. Discrepancies between scores on different subtests are often examined for their relation to psychiatric states. For ages sixteen years and over. For ages five to fifteen years the children's version is used: Wechsler Intelligence Scale for Children – W.I.S.C.

Wechsler Bellevue (Wechsler, 1947)
The fore-runner of the W.A.I.S.

Word Association (Woodworth and Schlosberg, 1954)

A list of words is read out to the patient who has to say the first word he thinks of as each word in the list is read out. Some standard word lists are available. Sometimes words are presented having particular significance for the patient. The reaction times and quality of association are taken as indicators of the emotional significance of stimulus words.

S.C.E.R.T., West Bengal
Library
Calcutta
25/3. B C.

Appendix 2
Tests Used in Clinical Psychology

The tests listed below are presented because they are ones which the clinical psychologist may expect to come across in his search for measures of behaviour and among which he will have to make choices. The list is not exhaustive and the classification is one of convenience. Each test has been listed only once whereas some of them could well have been listed under several categories.

Educational Attainments

Reading

Burt Reading Test
Gates–McKillop Reading Diagnostic Tests
Graded Word Reading Test (Schonell)
Graded Word Reading Test (Vernon)
Gray Oral Reading Test
Neale Analysis of Reading Ability

Arithmetic

Burt Arithmetic Tests
Schonell Arithmetic Tests

Intelligence and Abilities

Adult

A.H.5
Cattell Intelligence Tests
General Aptitude Test Battery
Illinois Test of Psycholinguistic Abilities
Kent Series of Emergency Scales

Mill Hill Vocabulary Scale
Nufferno Tests of Speed and Level
Porteus Maze Test
Primary Mental Abilities
Progressive Matrices – Standard
– Coloured
– Advanced
Shipley Institute of Living Scale for measuring intellectual impairment (Shipley Hartford)
Wechsler Adult Intelligence Scale
Wechsler–Bellevue Intelligence Scale

Children

Cattell Infant Intelligence Scale
Columbia Mental Maturity Scale
Full Range Picture Vocabulary Test
Goodenough–Harris Drawing Test (Goodenough Draw-a-Man Test revision)
Griffiths Mental Development Scale for testing babies from birth to two years
Merrill–Palmer Scale of Mental Tests
Peabody Picture Vocabulary Test
Stanford–Binet Intelligence Scale
Wechsler Intelligence Scale for Children

Performance Tests

Alexander Performance Scale
Arthur Point Scale of Performance Tests
Drever and Collins Performance Tests of Intelligence
Kohs Block Design Test

Memory

Benton Visual Retention Test
Memory-for-Designs Test
Wechsler Memory Scale
Williams Delayed Recall Test

Word Learning

Modified New Word Learning Test
Neologisms Learning Test
Paired Associate Learning Test
Synonym Learning Test

Speed

Reaction Time
Tapping rate

Perception

Archimedes Spiral After Effect
Bender Gestalt Test
Critical Flicker Fusion
Gottschaldt Embedded Figures
Marianne Frostig Developmental Test of Visual Perception
Minnesota Percepto-Diagnostic Test
Two Flash Threshold

Aphasia

Examining for Aphasia (Eisenson)

Concrete–Abstract Attitude

Chapman Figurative Literal Test
Diagnostic Performance Tests (Semeonoff–Trist)
Goldstein–Scheerer Tests of Abstract and Concrete Thinking
Grassi Block Substitution Test
Reitman's Pin-man Test
Shaw Test
Vigotsky Sorting Test

Schizophrenic Thinking Disorder

Chapman Overinclusion Test
Epstein's Test of Overinclusion
Grid Test (Bannister and Fransella)
Lovibond's Test of Overinclusion
Payne's Object Classification Test
Proverbs – Gorham's
- Benjamin's

Personality – Attitudes, Conflicts, Needs and Traits

Adult projective

Four Picture Test
Holtzman Inkblot Technique
House–Tree–Person Test
Kahn Test of Symbol Arrangement
Lowenfeld Mosaic Test
Machover Draw-a-Person Test
Make-a-Picture Story Test
Object Relations Technique
Rorschach Inkblots
Rosenzweig Picture-Frustration Study
Rotter Incomplete Sentences Blank
Symonds Picture Story Test
Szondi Test
Thematic Apperception Test
Tomkins–Horn Picture Arrangement Test

Adult questionnaires

Anxiety Scale Questionnaire
Bernreuter Personality Inventory
California Psychological Inventory
Dynamic Personality Inventory
Edwards Personal Preference Schedule
Edwards Social Desirability Scale
Eight-Parallel-Form Anxiety Battery
Eysenck Personality Inventory

Eysenck–Withers Personality Inventory
Guilford–Zimmerman Temperament Survey
Hostility and Direction of Hostility Questionnaire
Hysteroid–Obsessoid Questionnaire
Manifest Anxiety Scale
Marlowe–Crowne Social Desirability Scale
Maudsley Personality Inventory
Minnesota Multiphasic Personality Inventory
Motivation Analysis Test
Neuroticism Scale Questionnaire
Sixteen Personality Factor Questionnaire
Willoughby Personality Schedule

Adult other

Q sort
Repertory Grid
Semantic Differential
Word Association

Children projective

Blacky Pictures
Children's Apperception Test
Controlled Projection for Children
Pickford Projective Pictures
World Test

Children questionnaires

Children's Personality Questionnaire
High School Personality Questionnaire
Junior Eysenck Personality Questionnaire

Children other

Family Relations Indicator
Family Relations Test

Social Maturity

Progress Assessment Chart
Vineland Social Maturity Scale

Symptom Rating Scales and Check Lists

Beck's Depression Inventory
Bristol Social Adjustment Guides
Cornell Index
Elgin Prognostic Scale
Hamilton's Depression Inventory
Hospital Adjustment Scale
In-patient Multi-dimensional Psychiatric Scale
Mooney Problem Check List
Multi-dimensional Scale for Rating Psychiatric Patients
Psychotic Reaction Profile
Shapiro's Personal Questionnaire
Symptom Sign Inventory
Tavistock Self-Assessment Inventory
Wittenborn Psychiatric Rating Scales

Vocational Interests

Guilford–Zimmerman Interest Inventory
Kuder Preference Record
Strong Vocational Interest Blank
Study of Values
Thurstone Interest Schedule

Physiological Tests

Blood pressure
Body temperature
Electroencephalography
Electro-myography
Eye movement
Heart rate
Plethysmography
Pupillography

Respiration
Sedation threshold
Skin potential
Skin resistance

Miscellaneous

Harris Tests of Lateral Dominance
Ishihara Test for Colour Blindness
Level of Aspiration
Objective–Analytic Personality Test Batteries

References

Adams, M. B. (1964), '"Mental illness" or interpersonal behaviour' *Amer. Psychologist*, vol. 19, pp. 191–7.

Allport, F. H. (1955), *Theories of Perception and the Concept of Structure*, Wiley.

Allport, G. W. (1937), *Personality*, Constable.

Allport, G. W. (1961), *Pattern and Growth in Personality*, Holt, Rinehart & Winston.

American Psychological Association (1966 , *Standards for Educational and Psychological Tests and Manuals*, American Psychological Association, Washington, D.C.

Anastasi, A. (1961), *Psychological Testing*, Macmillan.

Bannister, D. (1965), 'The rationale and clinical relevance of repertory grid technique', *Brit. J. Psychiat*, vol. III, pp. 977–82.

Bender, L. (1938), *A Visual Motor Gestalt Test and its Clinical Use*, American Orthopsychiatric Association.

Bender, L. (1963), Foreword in *An Evaluation of the Bender-Gestalt Test*, A. Tolor and H. C. Schulberg, Charles C. Thomas.

Block, J. (1961 , *The Q-Sort Method in Personality Assessment and Psychiatric Research*, Charles C. Thomas.

Bramel, D. (1962), 'A dissonance theory approach to defensive projection', *J. abnorm. soc. Psychol.*, vol. 64, pp. 121–9.

Buros, O. K. (ed.) (1965), *The Sixth Mental Measurements Yearbook*, Gryphon.

Cattell, R. B. (1957), 'The conceptual and test distinction of neuroticism and anxiety', *J. clin. Psychol.*, vol. 13, pp. 221–33.

Cattell, R. B. (1965), *The Scientific Analysis of Personality*, Penguin.

Cattell, R. B., Baggaley, A. R., Checov, L., Cogan, E. A., Flint, D., Gruen, W., Husek, E., Meeland, T., Saunders, D. R., and Schiff, H. (1955), *Objective Analytic Personality Test Batteries*, Institute for Personality and Ability Testing, Illinois.

Cattell, R. B., and Eber, H. W. (1957), *Handbook for the Sixteen Personality Factor Questionnaire*, Institute for Personality and Ability Testing, Illinois.

Chapman, J., and McGhie, A. (1962), 'A comparative study of disordered attention in schizophrenia', *J. ment. Sci.*, vol. 108, pp. 487–500.

Chapman, J., and McGhie, A. (1963), 'An approach to the psychotherapy of cognitive dysfunction in schizophrenia', *Brit. J. med. Psychol.*, vol. 36, pp. 253–60.

Claridge, G. S., and Herrington, R. N. (1963), 'Excitation-inhibition and the theory of neurosis: a study of the sedation threshold', in *Experiments with Drugs*, Eysenck, H. J. (ed), Pergamon.

Cronbach, L. J. (1961) *Essentials of Psychological Testing*, Harper & Row.

Cronbach, L. J., and Gleser, G. C. (1965), *Psychological Tests and Personnel Decisions*, 2nd edn, University of Illinois Press.

Desai, M. M. (1965), 'Practical problems in the assessment of personality in the clinical field', *Brit. J. med. Psychol.*, vol. 38, pp. 231–40.

Epstein, S. (1953), 'Overinclusive thinking in a schizophrenic and a control group', *J. consult. Psychol.*, vol. 17, pp. 384–8.

Eysenck, H. J. (1947), *Dimensions of Personality*, Routledge & Kegan Paul.

Eysenck, H. J. (1952), *The Scientific Study of Personality*, Routledge & Kegan Paul.

Eysenck, H. J. (1960), 'Classification and the problem of diagnosis', in *Handbook of Abnormal Psychology*, Eysenck, H. J. (ed.), Pitman.

Eysenck, H. J. (1966), 'Personality and experimental psychology', *Bull. Brit. psychol. Soc.*, vol. 19, pp. 1–28.

Eysenck, H. J., and Eysenck, S. B. G. (1964), *Manual of the Eysenck Personality Inventory*, University of London Press.

Field, J. G. (1960), 'Two types of tables for use with Wechsler's intelligence scales', *J. clin. Psychol.*, vol. 16, pp. 1–7.

Finney, J. C. (1966), 'Programmed interpretation of M.M.P.I. and C.P.I.', *Arch. gen. Psychiat.*, vol. 15, pp. 75–81.

Foulds, G. A. (1961), 'Personality traits and neurotic symptoms and signs', *Brit. J. med. Psychol.*, vol. 34, pp. 263–70.

Foulds, G. A. (1962), 'A quantification of diagnostic differentiae', *J. ment. Sci.*, vol. 108, 389–405.

Foulds, G. A. (1965 , *Personality and Personal Illness*, Tavistock.

Freeman, T., and Gathercole, C. E. (1966), 'Perseveration – the clinical symptoms – in chronic schizophrenia and organic dementia', *Brit. J. Psychiat.*, vol. 112, pp. 27–32.

Freud, S. (1959), *Collected Papers*, Basic Books.

Furneaux, W. D. (1956), *Manual of Nufferno Speed Tests* and *Manual of Nufferno Level Tests*, Inst. of Psychiatry, London.

Furneaux, W. D. (1960), 'Intellectual abilities and problem solving behaviour', in *Handbook of Abnormal Psychology*, Eysenck, H. J. (ed.), Pitman.

Furneaux, W. D., and Lindahl, L. E. H. (1963), 'How valid are questionnaire validity studies?', *Bull. Brit. psychol. Soc.*, vol. 16, no. 51, p. A19.

Gathercole, C. E. (1966a), Unpublished case, Rainhill Hospital.

Gathercole, C. E. (1966b), 'I.Q. scores and the problem of classification', *Brit. J. Psychiat*, vol. 112, pp. 1181–2.

Gathercole, C. E. (1966c), 'What do clinical psychologists need to measure?' *J. Liverpool Psychiat. Soc.*, vol. 4, pp. 21–6.

Gedye, J. L. (1966), 'Experience with an automated psychological testing system designed for clinical use', *Bull. Brit. psychol. Soc.*, vol. 19, no. 63, p. A18.

Goldstein, K. (1939), *The Organism*, The American Book Co.

Goldstein, K. (1959), 'Functional disturbances in brain damage', in *American Handbook of Psychiatry*, Arieti, S. (ed.), Basic Books.

Goldstein, K., and Scheerer, M. (1941), 'Abstract and concrete behaviour', *Psychol. Monogr.*, vol. 53, no. 2.

Gough, H. G. (1957), *California Psychological Inventory*, Consulting Psychologists.

Graham, F. K., and Kendall, B. S. (1960), 'Memory for Designs Test: revised general manual', *Percept. motor Skills*, vol. 11, pp. 147–88.

Hamilton, V. (1964), 'Techniques and methods in psychological assessment', *Bull. Brit. psychol. Soc.*, vol. 17, pp. 27-36.

Hathaway, S. R., and McKinley, J. C. (1951), *Minnesota Multiphasic Personality Inventory Manual*, Psychological Corporation.

Heim, A. W. (1956), *Test A.H.5*, National Foundation for Educational Research, Slough.

Howells, J. G., and Lickorish, J. R. (1962), *The Family Relations Indicator*, National Foundation for Educational Research, Slough.

Inhelder, B., and Piaget, J. (1958), *The Growth of Logical Thinking from Childhood to Adolescence*, Routledge & Kegan Paul.

Jackson, J. H. (1938), *Selected Writings*, Staples.

Jaspers, K. (1913), *General Psychopathology*, English translation, 1963, Manchester University Press.

Jones, H. G. (1960), 'Applied abnormal psychology: the experimental approach', in *Handbook of Abnormal Psychology*, Eysenck, H. J. (ed.), Pitman.

Kendrick, D. C. (1965), 'Speed and learning in the diagnosis of diffuse brain damage in elderly subjects: a Bayesian statistical approach', *Brit. J. soc. clin. Psychol.*, vol. 4, pp. 141–8.

Kintz, B. L., Delprato, D. J., Mettee, D. R., Persons, C. E., and Schappe, R. H. (1965), 'The experimenter effect', *Psychol. Bull.*, vol. 3, pp. 223–32.

Klein, R., and Mayer-Gross, W. (1957), *The Clinical Examination of Patients with Organic Cerebral Disease*, Cassell.

Klopfer, B., and Davidson, H. H. (1962), *The Rorschach Technique: An Introductory Manual*, Harcourt, Brace & World.

Kornrich, M. (1965), *Psychological Test Modifications*, Charles C. Thomas.

Lader, M. H., and Wing, L. (1964), 'Habituation of the psychogalvanic reflex in patients with anxiety states and in normal subjects', *J. Neurol. Neurosurg. Psychiat.*, vol. 27, pp. 210–8.

Landis, C., and Hunt, W. A. (1939), *The Startle Pattern*, Rinehart.

Luria, A. R. (1966), *Higher Cortical Functions in Man*, Tavistock.

McGhie, A., and Chapman, J. (1961), 'Disorders of attention and perception in early schizophrenia', *Brit. J. med. Psychol.*, vol. 34, pp. 103–16.

McKellar, P. (1957), *Imagination and Thinking*, Cohen & West.

Marks, P. A., and Seeman, W. (1963), *The Actuarial Description of Abnormal Personality*, Williams & Wilkins.

Mayer-Gross, W., Slater, E., and Roth, M. (1960), *Clinical Psychiatry*, Cassell.

Meehl, P. E. (1954), *Clinical versus Statistical Prediction*, University of Minnesota Press.

Meehl, P. E., and Dahlstrom, W. G. (1960), 'Objective configural rules for discriminating psychotic from neurotic M.M.P.I. Profiles', *J. consult. Psychol.*, vol. 24, pp. 375–87.

Meehl, P. E., and Rosen, A. (1955), 'Antecedent probability and the efficiency of psychometric signs, patterns or cutting scores', *Psychol. Bull.*, vol. 52, pp. 194–216.

Meyer, V. (1960), 'Psychological effects of brain damage', in *Handbook of Abnormal Psychology*, Eysenck, H. J. (ed.), Pitman.

Moffat, J., and Levine, S. (1964), 'The sedation thresholds of psychoneurotic and alcoholic patients', *Brit. J. med. Psychol.*, vol. 37, pp. 313–17.

Murray, H. A. (1943), *Thematic Apperception Test*, Harvard University Press.

Nunnally, J. C. (1959), *Tests and Measurements*, McGraw-Hill.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957), *The Measurement of Meaning*, University of Illinois Press.

Pascal, G. R. and Suttell, B. J. (1951), *The Bender–Gestalt Test: its Quantification and Validity for Adults*, Grune & Stratton.

Payne, R. W. (1958), 'Diagnostic and personality testing in clinical psychology', *Amer. J. Psychiat.*, vol. 115, pp. 25–9.

Payne, R. W. (1962), 'An object classification test as a measure of overinclusive thinking in schizophrenic patients', *Brit. J. soc. clin. Psychol.*, vol. 1, pp. 213–21.

Payne, R. W., and Friedlander, D. 1962), 'A short battery of simple tests for measuring overinclusive thinking', *J. ment. Sci.*, vol. 108, pp. 362–7.

Payne, R. W., and Jones, H. G. (1957), 'Statistics for the investigation of individual cases', *J. clin. Psychol.*, vol. 13, pp. 115–21.

Penrose, L. S., and Raven, J. C. (1937), 'A new series of perceptual tests: preliminary communication', *Brit. J. med. Psychol.*, vol. 16, 97–104.

Piaget, J. (1929), *The Child's Conception of the World*, Harcourt, Brace.

Piercy, M. (1959), 'Testing for intellectual impairment – some comments on the tests and the testers', *J. ment. Sci.*, vol. 105, pp. 489–95.

Porteus, S. D. (1965), *Porteus Maze Test. Fifty Years' Application*, Pacific Books.

Rapaport, D. (1945), *Diagnostic Psychological Testing*, Year Book Pub.

Raven, J. C. (1958a), *Guide to using the Coloured Progressive Matrices*, Lewis.

Raven, J. C. (1958b), *Guide to using the Mill Hill Vocabulary Scale with the Progressive Matrices*, Lewis.

Raven, J. C. (1960), *Guide to the Standard Progressive Matrices*, Lewis.

Rodger, T. F., and Mowbray, R. M. (1961), 'Clinical method in psychological medicine', *J. med. Educ.*, vol. 36, pp. 167–72.

Rorschach, H. (1942), *Psychodiagnostics*, Huber.

Rosen, A. (1966), 'Stability of new M.M.P.I. scales and statistical procedures for evaluating changes and differences in psychiatric patients', *J. consult. Psychol.*, vol. 30, pp. 142–5.

Rosenthal, R. (1964), 'Experimenter outcome-orientation and the results of the psychological experiment', *Psychol. Bull.*, vol. 61, pp. 405–12.

Schafer, R. (1948), *The Clinical Application of Psychological Tests*, International Universities Press.

Schneider, K. (1958), *Psychopathic Personalities*, Cassell.

Shapiro, M. B. (1951), 'An experimental approach to diagnostic psychological testing', *J. ment. Sci.*, vol. 97, pp. 748–64.

Shapiro, M. B. (1961), 'A method of measuring psychological changes specific to the individual psychiatric patient', *Brit. J. med. Psychol.*, vol. 34, pp. 151–5.

Shipley, W. C. (1946), *Shipley Institute of Living Scale for Measuring Intellectual Impairment*, Shipley.

Slater, E. (1965), 'Diagnosis of hysteria', *Brit. med. J.*, vol. 1, pp. 1395–9.

Slater, P. (1965), 'The use of the repertory grid technique in the individual case', *Brit. J. Psychiat.*, vol. 111, pp. 965–75.

Standing Committee on Test Standards of the British Psychological Society (1965), 'Principles governing the employment of psychological tests and clinical instruments', *Bull. Brit. psychol. Soc.*, vol. 18, pp. 27–8.

Stevens, S. S. (1951), 'Mathematics, measurement and psychophysics', *Handbook of Experimental Psychology*, Stevens, S. S. (ed.), Wiley.

Sundberg, N. D., and Tyler, L. (1963), *Clinical Psychology*, Methuen.

Szasz, T. S. (1960), 'The myth of mental illness', *Amer. Psychologist*, vol. 15, pp. 113–18.

Taylor, J. A. (1953), 'A personality scale of manifest anxiety', *J. abnorm. soc. Psychol.*, vol. 48, pp. 285–90.

Thorndike, E. L., and Lorge, I. (1944), *The Teacher's Word Book of 30,000 Words*, Columbia University Press.

Tolor, A., and Schulberg, H. (1963), *An Evaluation of the Bender–Gestalt Test*, Charles C. Thomas.

United States Employment Service (1966), *General Aptitude Test Battery, U.S.E.S.*, Washington, 1966.

Venables, P. H., and Martin, I. (1967), 'Skin resistance and skin potential', in *Manual of Psycho-physiological Methods*, Venables, P. H., and Martin, I. (eds.), North Holland.

Walton, D. W., and Black, D. A. (1959), 'The predictive validity of a psychological test of brain damage', *J. ment. Sci.*, vol. 105, pp. 807–10.

Wechsler, D. (1947), *Wechsler Bellevue Intelligence Scale*, Psychological Corporation.

Wechsler, D. (1955), *Wechsler Adult Intelligence Scale*, Psychological Corporation.

Werner, H. (1948), *Comparative Psychology of Mental Development*, Follett.

White, R. W. (1952), 'What is tested by psychological tests?', in *Relation of Psychological Tests to Psychiatry*, Hoch, P. H., and Zubin, J. (eds.), Grune & Stratton.

Williams, M. (1965), *Mental Testing in Clinical Practice*, Pergamon.

Woodworth, R. S., and Schlosberg, H. (1954), *Experimental Psychology*, Methuen.

Zangwill, O. L. (1952), 'Discussion on the role of the psychologist in psychiatric practice', *Proc. Roy. Soc. Med.*, vol. 45, pp. 445–7.

Index

Abilities, tests of 83, 154–5
Abnormality
  concepts of 66–72
  of test discrepancies 48
Abstract attitude 156
Adams, M. B. 72
Administration of tests 109ff., 20, 89–90
Adults, tests for 154, 157, 158
Aggressive patients 115, 123–4
Alcoholism 34
Allport, F. H. 19, 20
Allport, G. W. 24, 62, 64
Amentia 35
American Psychological Association 84, 87
Anastasi, A. 87
Anorexia nervosa 39
Anxiety
  during testing 120–2, 130
  hysteria 74
  neurosis 69
  test of 113
Aphasia 117, 156
  see also Dysphasia
Arithmetic tests 83, 154
Arteriosclerosis, cerebral 88
Assessment procedures and tests
  A.H.5 43, 144, 154
  Archimedes spiral 21, 144, 156
  Bender's Visual Motor Gestalt Test 18, 85, 116, 118, 143, 145, 156
  Bernrenter Personality Inventory 91, 145, 157
  Body Sway Test of Suggestibility 22, 145
  California Psychological Inventory – C.P.I. 140, 145, 157
  Conservation Tests 116, 149
  Digit Copying Test 89
  Epstein's Test of Over-inclusive Thinking 119–20, 146, 157
  Eysenck Personality Inventory 63, 71, 79, 146, 158
  Family Relations Indicator – F.R.I. 18, 147, 158
  Galvanic Skin Response – G.S.R. 18, 21–2, 71, 112–3, 119, 122, 147, 160
  Goldstein–Scheerer Tests 18, 89, 147, 156
  Level of Aspiration 88, 147, 160
  Manifest Anxiety Scale 79, 86, 152, 158
  Memory for Designs Test 36, 99–101, 105–8, 148, 155
  Mill Hill Vocabulary Scale 36, 43, 66, 91, 120, 148, 155
  Minnesota Multiphasic Personality Inventory – M.M.P.I.
  automated 140
  construction and use 18, 67, 72, 138, 148, 158
  diagnosis of neurosis and psychosis 37, 45–6, 58
  Ego Strength Scale 47
  Modified New Word Learning Test – M.N.W.I.T. 87, 148, 156
  motor coordination 22

  Nufferno Speed and Level Tests 45, 148, 155
  Object Classification Test, 119, 157
  Objective – Analytic Personality Test Battery 85, 149, 160
  peg-board 123, 149
  Personal Questionnaire 64, 152, 159
  Porteus Mazes 20, 149, 155
  Progressive Matrices
    coloured 91, 146, 155
    standard 18, 36, 66, 91, 118–9, 127–8, 150, 155
  Q-sort 63, 150, 158
  reaction time – R.T. 140, 150, 156
  Repertory Grid Technique 63, 140, 150–51, 158
  Rorschach Ink Blots 18, 151, 157
  dangers in use 41, 72, 84–7, 104, 129
  subjective or objective 21–2
  Sedation Threshold 109, 113, 137–8, 151, 160
  semantic differential 63, 151, 158
  Shipley-Hartford-Scale 119, 152, 155
  Sixteen Personality Factor Questionnaire – 16 P.F. 73, 79, 91, 145, 158
  Sorting Test 84
  Story Recall 84
  tapping speed 21, 156
  Thematic Apperception Test – T.A.T. 22, 84–5, 91–2, 117, 128, 152, 157
  Wechsler Adult Intelligence Scale – W.A.I.S. 79, 85–6, 90–1, 98, 138, 152, 155
    Block Design 76–7, 117
    Comprehension 128
    Digit Span 118–9
    Object Assembly 119
    Similarities 88
    Vocabulary 56–7, 124
    V–P discrepancy 47–9
  Wechsler-Bellevue 47, 84, 153, 155
  Word Association 84, 153
Asthma 40
Aspirations 40
Attainments, educational 83, 85, 154
Attitudes, assessment of 40, 157
Attitudes of patients to testing 91, 110–1, 114, 120–6
Automation 132, 139–42
  administration of tests 130, 140–1
  analysis of test responses 140
  computers 141–2
  future of 142
  scoring 139

Bandwidth 18, 86
Bannister, D. 150
Base-rates 93ff, 26, 49, 134, 138, 142
  detecting 102–5
Behaviour therapy 40, 92
Bender, L. 18, 116, 143, 145
Black, D. A. 87, 148
Block, J. 150
Brain disorder 93ff.
  assessment of 18, 48, 51, 56, 65–9, 79, 89, 113, 117, 123–5
  referrals 30, 33–4, 80–81, 88
Bramel, D. 122
British Psychological Society 83
Buros, O. K. 87, 144, 145

Carbon monoxide poisoning 34
Catastrophic reaction 123–4
Cattell, R. B. 18, 22, 62–3, 71, 73–4, 84–5, 145, 149

Causes 23, 64–7
  see also Diagnosis, implications for etiology
Chapman, J. 116, 125, 135, 146
Check-lists 84, 159
Children 33, 74–5, 85, 125
  tests for 155, 158
Claridge, G. S. 109, 138, 151
Cognitive-impairment 34, 48, 55, 88
  see also Brain disorder
Compulsive gambling 39
Computer 50, 140–2
Concrete attitude 156
Conditioning, verbal 127–8
Confabulation 34
Conflicts 28, 30, 40, 157
Continuity 77–82
Conversion symptoms 77
Cooperation of patients in testing 109, 114–5
Court cases 36
Cross-validation 46, 107
Cronbach, L. J. 27, 86, 87
Cyclical-vomiting 39

Dahlstrom, W. G. 37, 45, 58, 140
Davidson, H. H. 151
Decisions about patients 17–9, 25–8, 42–3, 50, 74, 93
Defective, mental 32, 65, 68–70, 79
  see also Subnormal
Defence mechanisms 30, 40
Delinquent 79
Delirium 69
Delusions 77, 82, 124
Dementia 33, 35, 75, 80–81, 116, 123
Denotation 20
Desai, M. M. 27
Developmental dimension 74–5, 78
Depression 33, 51, 81, 99, 117, 122
Diagnosis 29–30, 37–9, 76, 133
  implications for
    description 37–8, 134, 138
    etiology 37–8, 67, 138
    prognosis 37–8, 134, 138
    treatment 37–8, 134, 138
Disease, concept of 66–72
Dissociation 78
Dissolution 75
Distractibility 124–5
  tests of 116, 119, 135, 146
Dypsomania 39
Dysphasia 34–5
  see also Aphasia
Dysthymia 137

Early data returns 128
Ebbinghaus, H. 137
Eber, H. W. 74, 145
Elation 122
Encouragement in testing 118–9
Epstein, S. 119, 146
Evolution 75
Experimenter effect 126–32, 134
Eysenck, H. J. 63, 68, 71–3, 145–7
Eysenck, S. B. G. 71, 146

Factor analysis 72–6
Fatigue 18, 125
Fidelity 18, 86
Field, J. G. 47–9
Finney, J. C. 140–42
Fits 34
Follow-up 53, 87, 93, 102, 134
Foulds, G. A. 30, 71–3, 77, 79–80
Freeman, T. 119, 123
Freud, S. 75, 137
Friedlander, D. 30, 81
Frustration tolerance 77
Furneaux, W. D. 45, 76, 119, 148

Ganser symptoms 33
Gathercole, C. E. 55, 98, 119, 123

Gedge, J. L. 132, 141
General Paralysis of the Insane – G.P.I. 34, 67
Genotypes 64–6, 71, 77
Gleser, G. C. 27, 86
Goldstein, K. 18, 75, 81, 89, 124, 147
Gough, H. G. 145
Graham, F. K. 99–101, 105–7, 148

Habituation of G.S.R. 113, 119
Hallucinations 77–8
Hamilton, V. 57
Hathaway, S. R. 18, 67, 148
Hein, A. W. 144
Herrington, R. N. 109, 138, 151
Homosexual 75
Howells, J. G. 18, 147
Hunt, W. A. 122
Hypnagogic 78
Hysteria and hysterical 33, 69, 72, 93, 103, 138

Idiographic 27, 62–4, 75, 137
Imagery 78
Indecent exposure 39
Individuation 78
Inference 44ff.
  clinical 22, 32, 50–4, 86, 89–90, 131, 133
  from experimental study 44, 54–60, 86, 89–90, 131
  from tests 21, 23, 38, 105
  psychiatrists 18
  statistical 22, 44–50, 86, 88–90, 131
  subjective 20, 33
Inhelder, B. 116
Instrumentation factors 84
Intelligence 76, 78–9
  assessment of 45, 63
  referrals 32, 35–6
  tests 154–5
Interests 40–41, 46, 84, 159

Jackson, J. H. 75
Jaspers, K. 68
Jones, H. G. 26, 39, 46, 88, 140
Jung, C. G. 18

Kendall, B. S. 99–101, 105–107, 148
Kendrick, D. C. 89
Kintz, B. L. 129
Klein, R. 18
Klopfer, B. 151
Kornrich, M. 30
Kraepelin, E. 18
Kretschmer, E. 18

Lader, M. H. 113
Landis, C. 122
Learning 24, 33
  tests of 18, 156
Leucotomy 34, 71
Level of aspiration 77, 88
Levine, S. 138
Lewin, K. 64
Lickorish, J. R. 18, 147
Lindahl, L. E. H. 119
Localization of brain disorder 35
Localization of brain disorder 3
Lorge, I. 56
L.S.D. 124
Luria, A. R. 81

McGhie, A. 116, 125, 135
McKellar, P. 78
McKinley, J. C. 18, 67, 148
Malingering 86
Mania 64
Marks, P. A. 45
Martin, I. 147
Mass action, law of 80
Maturity 86, 159
Mayer-Gross, W. 18, 22, 69
Measurement 19, 22–24, 48, 71, 135
Mechanisms of behaviour 30, 75–6, 78

Meehl, P. E. 37, 39, 44, 45, 50, 58, 59, 93, 96, 104, 140
Memory 18, 24, 33–4, 85, 88, 138, 155
Meyer, V. 35
Microcephaly 68
Moffat, J. 138
Motivation 24, 58
Mowbray, R. M. 51
Multiple choice tests 21, 127
Murray, H. A. 152
Myxoedema 34

Needs 40, 157
Neurasthenia 69
Neurosis 32, 69–73, 78, 81, 95
  assessment of 39, 45–6, 58, 63, 138
Neuroticism 63, 72–3
Nomothetic 27, 62–4, 72, 136, 138
Normality 66
Norms 20, 23, 26, 46, 55, 62, 89–90
Nunnally, J. C. 87

Objectivity 19–23, 27, 30, 32, 51, 58, 83, 86–7
Obsessional 55, 72, 80
Open-ended tests 21, 89, 128
Orientation of psychiatrist 30
  of psychologist 27–8, 30, 54, 59–60, 89
Osgood, C. E. 63, 151
Overinclusive thinking 119

Paraphrenia 82
  affect-laden 122
Pavlov, I. P. 137
Payne, R. W. 30, 37–8, 46, 81, 119, 140
Penrose, L. S. 18
Perceptual tests 156
Performance tests 155
Perseveration 75, 119–20, 123–5
Persistence 40, 76
Personality
  abnormal 69
  dimensions 63, 67, 72–5, 78–82, 86
  structure and function 25, 27–8, 72–7
  tests 40, 157
  traits 22, 63, 67, 71–2, 80, 157
Phenotypes 64–6, 71, 77
Phenylketonuria 68
Phobia 32
Physiological tests 159–60
Piaget, J. 75, 116, 149
Piercy, M. 26
Porphyria 69
Porteus, S. D. 20, 14, 9
Predisposition 69–71
Problem solving 77
Profile, tests 46, 74
Progression 75
Projective tests 21–2, 32, 41, 157–8
Prompting in testing 117–9
Proverbs 36, 157
Psoriasis 39
Psychiatric examination 17–19
Psychiatrists
  collaboration with psychologists 18, 29, 31, 41–2
  training 19
  use of tests 18
Psychology
  application to psychitary 19, 24, 28
  as applied science 24–6, 135
Psychomotor tasks 123
Psychopaths 69, 109
Psychosis 64, 72, 78, 81, 113
  assessment of 45–6, 58, 124
  paranoid 82
Psychosomatic 39, 88
Psychotherapy 30, 35, 43, 47, 91–2, 120
Psychotomimetics 78
  see also L.S.D.

Quantitative and qualitative 23, 62, 77–82, 89
Questionnaires 21–2, 47, 84, 91, 119, 139, 157–8

Rapaport, D. 76, 84
Rapport 112
Rating scales 159
Rats, maze bright 127
Raven, J. C. 18, 36, 146, 148, 150
Reading tests 154
Reassurance in testing 115, 118–9, 121–2
Referrals 29ff. 54, 96, 103, 107, 111
Regression 75
Reliability of tests 21, 24, 26, 47, 87
Repression 78
Research 133ff. 19, 24–5, 33, 50, 54, 60, 66, 109, 113, 128
Retardation 117
Rodger, T. F. 51
Rorschach, H. 18
Rosen, A. 46–7, 93, 96, 104
Rosenthal, R. 126, 129
Roth, M. 22, 69

Schafer, R. 84
Scheerer, M. 18, 89, 147
Schizophrenia 33, 67, 113
  assessment 45, 47, 51, 74–5, 115–6, 121, 124, 134–6
  hebephrenic 115
  paranoid 82
  simple 35
Schlosberg, H. 150, 153
School refusal 39
Schneider, K. 68–9, 71
Schulberg, H. 76, 143
Seeman, W. 45
Self-administered tests 130
Self-concept 40, 63
Shapiro, M. B. 26, 54, 64, 152
Shipley, W. C. 152
Slater, E. 22, 69, 93, 103
Slater, P. 140
Speed 33
  and intelligence 76
  tests of 45, 156
Standardization of tests 20, 26, 32, 44, 48, 49, 87, 89–90, 134, 138
Stevens, S. S. 23
Stress 40, 70–1
Subjectivity 21, 27, 44, 58, 133
Subnormal 35, 98
Sundberg, N. D. 87
Szasz, T. S. 68, 71

Taylor, J. A. 152
Temperament 58, 66
Tension 17
Testers
  expectations 126–32
  personality 125–6
Testing technique 109ff.
Tests
  see also Assessment procedures and tests
  batteries 48, 84–7
  changes in scores 46–7
  cutting scores 93ff.
  defined 84–5
  descriptions of selected 144ff.
  development 19, 33
  efficiency 96–101
  error 26, 47, 66, 105
  evaluation of 87–8, 144
  limitations 26–8, 37–9, 41, 63, 85, 87–8, 133–4
  list of 154ff.
  misclassifications 101–2
  modifications 30, 90, 117
  need for 17–9
  order of administration 119–20

Tests – *continued*
  probabilities and scores 93ff., 45, 47, 51, 138
  scoring 20–21
  see also Automation
  selection 83ff., 48
  used by psychiatrists 18
  usefulness 93ff., 27, 66, 87
Theory 25, 27–8, 61, 76, 142–3
Thinking disorder 81, 82, 103, 134, 157
Thorndike, E. L. 56
Thought blocking 64
Thyrotoxicosis 69
Tolor, A. 76, 143
Tyler, L. 87

Uncertainty, tolerance of 25–8
Untestable patients 111

Validity and validation 93ff., 24, 26–7, 32, 37–8, 87, 90, 110, 133–4
  of clinical judgement 52–3, 133
Venables, P. H. 147
Vocabulary tests 56–7

Walton, D. W. 87, 148
Wechsler, D. 76, 117, 152, 153
Werner, H. 75
White, R. W. 77
Williams, M. 38, 76
Wing, L. 113
Woodworth, R. S. 150, 153
Word frequency 56–7
Wundt, W. M. 18

Zangwill, O. L. 53

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

Penguin science of behaviour

132·075
GAT

Assessment in clinical psychology offers an introduction and a framework to some of the main issues of assessment. Throughout this account of the practical contribution of the clinical psychologist, the author stresses the need for and the advantages of psychological testing, but he also draws attention to many shortcomings.

The author describes some of the problems with which psychiatrists and others present the clinical psychologist. These problems include diagnostic assessment, evaluation of the effects of treatment, quantification of personality characteristics, determination of intellectual level, speed and persistence and their appropriateness for a particular treatment or occupation, and many others.

This clear, unillusioned view gives a special authority to Mr Gathercole's survey. Two appendixes, describing and classifying many of the major tests used in clinical psychology, make the book also a handy reference source.

C. E. Gathercole is Senior Psychologist at the Benedict Clinic, Rainhill Hospital near Liverpool.

United King p 6
Australia $1.00
New Zealand $0.90
South Africa R0.75
Canada $1.25

64

Cover design by
Martin Bassett

Published by Penguin Education